PCT







INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6: C12N 15/31, C07K 14/22, 16/12, C12Q 1/68, A61K 39/095, G01N 33/50

A2

(11) International Publication Number:

WO 99/57280

(43) International Publication Date:

11 November 1999 (11.11.99)

(21) International Application Number:

PCT/US99/09346

(22) International Filing Date:

30 April 1999 (30.04.99)

(30) Priority Data:

morny Data:		
60/083,758	1 May 1998 (01.05.98)	US
60/094,869	31 July 1998 (31.07.98)	US
60/098,994	2 September 1998 (02.09.98)	US
60/099,062	2 September 1998 (02.09.98)	US
60/103,749	9 October 1998 (09.10.98)	US
60/103,794	9 October 1998 (09.10.98)	US
60/103,796	9 October 1998 (09.10.98)	US
60/121,528	25 February 1999 (25.02.99)	US

(71) Applicants (for all designated States except US): CHIRON CORPORATION [US/US]; 4560 Horton Street, Emeryville, CA 94608 (US). THE INSTITUTE FOR GENOMIC RE-SEARCH [US/US]; 9212 Medical Center Drive, Rockville, MD 20850 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): FRASER, Claire [US/US]; Rockville, MD (US). GALEOTTI, Cesira [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). GRANDI, Guido [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). HICKEY, Erin [US/US]; Gaithersburg, MD

(US). MASIGNANI, Vega [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). MORA, Marirosa [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). PETERSEN, Jeremy [US/US]; Arlington, VA (US). PIZZA, Mariagratia [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). RAPPUOLI, Rino [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). RATTI, Giulio [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). SCALATO, Enzo [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). SCARSELLI, Maria [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). TETTELIN, Herve [US/US]; Gaithersburg, MD (US). VENTER, J., Craig [US/US]; Rockville, MD (US).

(74) Agent: HARBIN, Alisa, A.; Chiron Corporation, Intellectual Property - R440, P.O. Box 8097, Emeryville, CA 94662-8097 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: NEISSERIA MENINGITIDIS ANTIGENS AND COMPOSITIONS

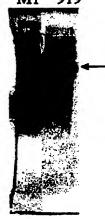
(57) Abstract

The invention provides proteins from Neisseria meningitidis, including the amino acid sequences and the corresponding nucleotide sequences. The proteins are predicted to be useful antigens for vaccines and/or diagnostics.

919 (46 kDa)

A) PURIFICATION

MI 919



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

NEISSERIA MENINGITIDIS ANTIGENS AND COMPOSITIONS

This application is a continuation-in-part of the following U.S. Provisional Patent applications, from each of which priority is claimed, and each of which is incorporated by reference in its entirety: 60/083,758 (filed May 1, 1998); 60/094,869 (filed July 31, 1998); 60/098,994 (filed September 2, 1998); 60/099,062 (filed September 2, 1998); 60/103,749 (filed October 9, 1998); 60/103,794 (filed October 9, 1998); 60/103,796 (filed October 9, 1998); and 60/121,528 (filed February 25, 1999).

This invention relates to antigens from the bacterial species: Neisseria meningitidis and Neisseria gonorrhoeae.

BACKGROUND

Neisseria meningitidis is a non-motile, gram negative diplococcus human pathogen. It colonizes the pharynx, causing meningitis and, occasionally, septicaemia in the absence of meningitis. It is closely related to N. gonorrhoea, although one feature that clearly differentiates meningococcus from gonococcus is the presence of a polysaccharide capsule that is present in all pathogenic meningococci.

N. meningitidis causes both endemic and epidemic disease. In the United States the attack rate is 0.6-1 per 100,000 persons per year, and it can be much greater during outbreaks. (see Lieberman et al. (1996) Safety and Immunogenicity of a Serogroups A/C Neisseria meningitidis Oligosaccharide-Protein Conjugate Vaccine in Young Children. JAMA 275(19):1499-1503; Schuchat et al (1997) Bacterial Meningitis in the United States in 1995. N Engl J Med 337(14):970-976). In developing countries, endemic disease rates are much higher and during epidemics incidence rates can reach 500 cases per 100,000 persons per year. Mortality is extremely high, at 10-20% in the United States, and much higher in developing countries. Following the introduction of the conjugate vaccine against Haemophilus influenzae, N. meningitidis is the major cause of bacterial meningitis at all ages in the United States (Schuchat et al (1997) supra).

Based on the organism's capsular polysaccharide, 12 serogroups of N. meningitidis have been identified. Group A is the pathogen most often implicated in epidemic disease in sub-Saharan Africa. Serogroups B and C are responsible for the vast majority of cases in the

United States and in most developed countries. Serogroups W135 and Y are responsible for the rest of the cases in the United States and developed countries. The meningococcal vaccine currently in use is a tetravalent polysaccharide vaccine composed of serogroups A, C, Y and W135. Although efficacious in adolescents and adults, it induces a poor immune response and short duration of protection, and cannot be used in infants [eg. Morbidity and Mortality weekly report, Vol.46, No. RR-5 (1997)]. This is because polysaccharides are T-cell independent antigens that induce a weak immune response that cannot be boosted by repeated immunization. Following the success of the vaccination against *H.influenzae*, conjugate vaccines against serogroups A and C have been developed and are at the final stage of clinical testing (Zollinger WD "New and Improved Vaccines Against Meningococcal Disease". In: New Generation Vaccines, supra, pp. 469-488; Lieberman et al (1996) supra; Costantino et al (1992) Development and phase I clinical testing of a conjugate vaccine against meningococcus A and C. Vaccine 10:691-698).

Meningococcus B (menB) remains a problem, however. This serotype currently is responsible for approximately 50% of total meningitis in the United States, Europe, and South America. The polysaccharide approach cannot be used because the menB capsular polysaccharide is a polymer of $\alpha(2-8)$ -linked N-acetyl neuraminic acid that is also present in mammalian tissue. This results in tolerance to the antigen; indeed, if an immune response were elicited, it would be anti-self, and therefore undesirable. In order to avoid induction of autoimmunity and to induce a protective immune response, the capsular polysaccharide has, for instance, been chemically modified substituting the N-acetyl groups with N-propionyl groups, leaving the specific antigenicity unaltered (Romero & Outschoorn (1994) Current status of Meningococcal group B vaccine candidates: capsular or non-capsular? Clin Microbiol Rev 7(4):559-575).

Alternative approaches to menB vaccines have used complex mixtures of outer membrane proteins (OMPs), containing either the OMPs alone, or OMPs enriched in porins, or deleted of the class 4 OMPs that are believed to induce antibodies that block bactericidal activity. This approach produces vaccines that are not well characterized. They are able to protect against the homologous strain, but are not effective at large where there are many antigenic variants of the outer membrane proteins. To overcome the antigenic variability, multivalent vaccines containing up to nine different porins have been constructed (eg. Poolman JT (1992) Development of a meningococcal vaccine. Infect. Agents Dis. 4:13-28).

·

Additional proteins to be used in outer membrane vaccines have been the opa and opc proteins, but none of these approaches have been able to overcome the antigenic variability (eg. Ala'Aldeen & Borriello (1996) The meningococcal transferrin-binding proteins 1 and 2 are both surface exposed and generate bactericidal antibodies capable of killing homologous and heterologous strains. Vaccine 14(1):49-53).

A certain amount of sequence data is available for meningococcal and gonoccocal genes and proteins (e.g. EP-A-0467714, WO96/29412), but this is by no means complete. Other men B proteins may include those listed in WO 97/28273, WO 96/29412, WO 95/03413, US 5,439,808, and US 5,879,686.

The provision of further sequences could provide an opportunity to identify secreted or surface-exposed proteins that are presumed targets for the immune system and which are not antigenically variable. For instance, some of the identified proteins could be components of efficacious vaccines against meningococcus B, some could be components of vaccines against all meningococcal serotypes, and others could be components of vaccines against all pathogenic Neisseriae including Neisseria meningitidis or Neisseria gonorrhoeae. Those sequences specific to N. meningitidis or N. gonorrhoeae that are more highly conserved are further preferred sequences.

It is thus an object of the invention is to provide Neisserial DNA sequences which encode proteins that are antigenic or immunogenic.

BRIEF DESCRIPTION OF THE DRAWINGS

- Fig. 1 illustrates the products of protein expression and purification of the predicted ORF 919 as cloned and expressed in *E. coli*.
- Fig. 2 illustrates the products of protein expression and purification of the predicted ORF 279 as cloned and expressed in *E. coli*.
- Fig. 3 illustrates the products of protein expression and purification of the predicted ORF 576-1 as cloned and expressed in *E. coli*.
- Fig. 4 illustrates the products of protein expression and purification of the predicted ORF 519-1 as cloned and expressed in *E. coli*.
- Fig. 5 illustrates the products of protein expression and purification of the predicted ORF 121-1 as cloned and expressed in *E. coli*.

Fig. 6 illustrates the products of protein expression and purification of the predicted ORF 128-1 as cloned and expressed in *E. coli*.

Fig. 7 illustrates the products of protein expression and purification of the predicted ORF 206 as cloned and expressed in *E. coli*.

Fig. 8 illustrates the products of protein expression and purification of the predicted ORF 287 as cloned and expressed in *E. coli*.

Fig. 9 illustrates the products of protein expression and purification of the predicted ORF 406 as cloned and expressed in *E. coli*.

Fig. 10 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 919 as cloned and expressed in *E. coli*.

Fig. 11 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 279 as cloned and expressed in *E. coli*.

Fig. 12 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 576-1 as cloned and expressed in *E. coli*.

Fig. 13 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 519-1 as cloned and expressed in *E. coli*.

Fig. 14 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 121-1 as cloned and expressed in *E. coli*.

Fig. 15 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 128-1 as cloned and expressed in *E. coli*.

Fig. 16 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 206 as cloned and expressed in *E. coli*.

Fig. 17 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 287 as cloned and expressed in *E. coli*.

Fig. 18 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 406 as cloned and expressed in *E. coli*.

Fig. 19 shows an alignment comparison of amino acid sequences for ORF 225 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

Fig. 20 shows an alignment comparison of amino acid sequences for ORF 235 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

Fig. 21 shows an alignment comparison of amino acid sequences for ORF 287 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

Fig. 22 shows an alignment comparison of amino acid sequences for ORF 519 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

Fig. 23 shows an alignment comparison of amino acid sequences for ORF 919 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

THE INVENTION

The invention provides proteins comprising the *N. meningitidis* amino acid sequences and *N. gonorrhoeae* amino acid sequences disclosed in the examples.

It also provides proteins comprising sequences homologous (i.e., those having sequence identity) to the *N. meningitidis* amino acid sequences disclosed in the examples. Depending on the particular sequence, the degree of homology (sequence identity) is preferably greater than 50% (eg. 60%, 70%, 80%, 90%, 95%, 99% or more). These proteins include mutants and allelic variants of the sequences disclosed in the examples. Typically, 50% identity or more between two proteins is considered to be an indication of functional equivalence. Identity between proteins is preferably determined by the Smith-Waterman

homology search algorithm as implemented in MPSRCH program (Oxford Molecular) using an affine gap search with parameters:gap penalty 12, gap extension penalty 1.

The invention further provides proteins comprising fragments of the N. meningitidis amino acid sequences and N. gonorrhoeae amino acid sequences disclosed in the examples. The fragments should comprise at least n consecutive amino acids from the sequences and, depending on the particular sequence, n is 7 or more (eg. 8, 10, 12, 14, 16, 18, 20 or more). Preferably the fragments comprise an epitope from the sequence.

The proteins of the invention can, of course, be prepared by various means (eg. recombinant expression, purification from cell culture, chemical synthesis etc.) and in various forms (eg. native, fusions etc.). They are preferably prepared in substantially pure or isolated form (ie. substantially free from other N. meningitidis or N. gonorrhoeae host cell proteins)

According to a further aspect, the invention provides antibodies which bind to these proteins. These may be polyclonal or monoclonal and may be produced by any suitable means.

According to a further aspect, the invention provides nucleic acid comprising the N. meningitidis nucleotide sequences and N. gonorrhoeae nucleotide sequences disclosed in the examples.

According to a further aspect, the invention comprises nucleic acids having sequence identity of greater than 50% (e.g., 60%, 70%, 80%, 90%, 95%, 99% or more) to the nucleic acid sequences herein. Sequence identity is determined as above-discussed.

According to a further aspect, the invention comprises nucleic acid that hybridizes to the sequences provided herein. Conditions for hybridization are set forth herein.

Nucleic acid comprising fragments of these sequences are also provided. These should comprise at least n consecutive nucleotides from the N. meningitidis sequences or N. gonorrhoeae sequences and depending on the particular sequence, n is 10 or more (eg 12, 14, 15, 18, 20, 25, 30, 35, 40 or more).

According to a further aspect, the invention provides nucleic acid encoding the proteins and protein fragments of the invention.

It should also be appreciated that the invention provides nucleic acid comprising sequences complementary to those described above (eg. for antisense or probing purposes).

Nucleic acid according to the invention can, of course, be prepared in many ways (eg. by chemical synthesis, in part or in whole, from genomic or cDNA libraries, from the

organism itself etc.) and can take various forms (eg. single stranded, double stranded, vectors, probes etc.).

In addition, the term "nucleic acid" includes DNA and RNA, and also their analogues, such as those containing modified backbones, and also protein nucleic acids (PNA) etc.

According to a further aspect, the invention provides vectors comprising nucleotide sequences of the invention (eg. expression vectors) and host cells transformed with such vectors.

According to a further aspect, the invention provides compositions comprising protein, antibody, and/or nucleic acid according to the invention. These compositions may be suitable as vaccines, for instance, or as diagnostic reagents or as immunogenic compositions.

The invention also provides nucleic acid, protein, or antibody according to the invention for use as medicaments (eg. as vaccines) or as diagnostic reagents. It also provides the use of nucleic acid, protein, or antibody according to the invention in the manufacture of (I) a medicament for treating or preventing infection due to Neisserial bacteria (ii) a diagnostic reagent for detecting the presence of Neisserial bacteria or of antibodies raised against Neisserial bacteria or (iii) for raising antibodies. Said Neisserial bacteria may be any species or strain (such as N. gonorrhoeae) but are preferably N. meningitidis, especially strain B or strain C.

The invention also provides a method of treating a patient, comprising administering to the patient a therapeutically effective amount of nucleic acid, protein, and/or antibody according to the invention.

According to further aspects, the invention provides various processes.

A process for producing proteins of the invention is provided, comprising the step of culturing a host cell according to the invention under conditions which induce protein expression.

A process for detecting polynucleotides of the invention is provided, comprising the steps of: (a) contacting a nucleic probe according to the invention with a biological sample under hybridizing conditions to form duplexes; and (b) detecting said duplexes.

A process for detecting proteins of the invention is provided, comprising the steps of:
(a) contacting an antibody according to the invention with a biological sample under conditions suitable for the formation of an antibody-antigen complexes; and (b) detecting said complexes.

Having now generally described the invention, the same will be more readily understood through reference to the following examples which are provided by way of illustration, and are not intended to be limiting of the present invention, unless specified.

Methodology - Summary of standard procedures and techniques. General

This invention provides Neisseria meningitidis menB nucleotide sequences, amino acid sequences encoded therein. With these disclosed sequences, nucleic acid probe assays and expression cassettes and vectors can be produced. The expression vectors can be transformed into host cells to produce proteins. The purified or isolated polypeptides (which may also be chemically synthesized) can be used to produce antibodies to detect menB proteins. Also, the host cells or extracts can be utilized for biological assays to isolate agonists or antagonists. In addition, with these sequences one can search to identify open reading frames and identify amino acid sequences. The proteins may also be used in immunogenic compositions, antigenic compositions and as vaccine components.

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature e.g., Sambrook Molecular Cloning; A Laboratory Manual, Second Edition (1989); DNA Cloning, Volumes I and ii (D.N Glover ed. 1985); Oligonucleotide Synthesis (M.J. Gait ed, 1984); Nucleic Acid Hybridization (B.D. Hames & S.J. Higgins eds. 1984); Transcription and Translation (B.D. Hames & S.J. Higgins eds. 1984); Animal Cell Culture (R.I. Freshney ed. 1986); Immobilized Cells and Enzymes (IRL Press, 1986); B. Perbal, A Practical Guide to Molecular Cloning (1984); the Methods in Enzymology series (Academic Press, Inc.), especially volumes 154 & 155; Gene Transfer Vectors for Mammalian Cells (J.H. Miller and M.P. Calos eds. 1987, Cold Spring Harbor Laboratory); Mayer and Walker, eds. (1987), Immunochemical Methods in Cell and Molecular Biology (Academic Press, London); Scopes, (1987) Protein Purification: Principles and Practice, Second Edition (Springer-Verlag, N.Y.), and Handbook of Experimental Immunology, Volumes I-IV (D.M. Weir and C.C. Blackwell eds 1986).

Standard abbreviations for nucleotides and amino acids are used in this specification.

All publications, patents, and patent applications cited herein are incorporated $\underline{\text{in full}}$ by reference.

Expression systems

The Neisseria menB nucleotide sequences can be expressed in a variety of different expression systems; for example those used with mammalian cells, plant cells, baculoviruses, bacteria, and yeast.

i. Mammalian Systems

Mammalian expression systems are known in the art. A mammalian promoter is any DNA sequence capable of binding mammalian RNA polymerase and initiating the downstream (3') transcription of a coding sequence (e.g., structural gene) into mRNA. A promoter will have a transcription initiating region, which is usually placed proximal to the 5' end of the coding sequence, and a TATA box, usually located 25-30 base pairs (bp) upstream of the transcription initiation site. The TATA box is thought to direct RNA polymerase II to begin RNA synthesis at the correct site. A mammalian promoter will also contain an upstream promoter element, usually located within 100 to 200 bp upstream of the TATA box. An upstream promoter element determines the rate at which transcription is initiated and can act in either orientation (Sambrook et al. (1989) "Expression of Cloned Genes in Mammalian ** Cells." In Molecular Cloning: A Laboratory Manual, 2nd ed.).

Mammalian viral genes are often highly expressed and have a broad host range; therefore sequences encoding mammalian viral genes provide particularly useful promoter sequences. Examples include the SV40 early promoter, mouse mammary tumor virus LTR promoter, adenovirus major late promoter (Ad MLP), and herpes simplex virus promoter. In addition, sequences derived from non-viral genes, such as the murine metallothionein gene, also provide useful promoter sequences. Expression may be either constitutive or regulated (inducible). Depending on the promoter selected, many promotes may be inducible using known substrates, such as the use of the mouse mammary tumor virus (MMTV) promoter with the glucocorticoid responsive element (GRE) that is induced by glucocorticoid in hormone-responsive transformed cells (see for example, U.S. Patent 5,783,681).

The presence of an enhancer element (enhancer), combined with the promoter elements described above, will usually increase expression levels. An enhancer is a

regulatory DNA sequence that can stimulate transcription up to 1000-fold when linked to homologous or heterologous promoters, with synthesis beginning at the normal RNA start site. Enhancers are also active when they are placed upstream or downstream from the transcription initiation site, in either normal or flipped orientation, or at a distance of more than 1000 nucleotides from the promoter (Maniatis et al. (1987) Science 236:1237; Alberts et al. (1989) Molecular Biology of the Cell, 2nd ed.). Enhancer elements derived from viruses may be particularly useful, because they usually have a broader host range. Examples include the SV40 early gene enhancer (Dijkema et al (1985) EMBO J. 4:761) and the enhancer/promoters derived from the long terminal repeat (LTR) of the Rous Sarcoma Virus (Gorman et al. (1982b) Proc. Natl. Acad. Sci. 79:6777) and from human cytomegalovirus (Boshart et al. (1985) Cell 41:521). Additionally, some enhancers are regulatable and become active only in the presence of an inducer, such as a hormone or metal ion (Sassone-Corsi and Borelli (1986) Trends Genet. 2:215; Maniatis et al. (1987) Science 236:1237).

A DNA molecule may be expressed intracellularly in mammalian cells. A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus of the recombinant protein will always be a methionine, which is encoded by the ATG start codon. If desired, the N-terminus may be cleaved from the protein by *in vitro* incubation with cyanogen bromide.

Alternatively, foreign proteins can also be secreted from the cell into the growth media by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provides for secretion of the foreign protein in mammalian cells. Preferably, there are processing sites encoded between the leader fragment and the foreign gene that can be cleaved either *in vivo* or *in vitro*. The leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell. The adenovirus tripartite leader is an example of a leader sequence that provides for secretion of a foreign protein in mammalian cells.

Usually, transcription termination and polyadenylation sequences recognized by mammalian cells are regulatory regions located 3' to the translation stop codon and thus, together with the promoter elements, flank the coding sequence. The 3' terminus of the mature mRNA is formed by site-specific post-transcriptional cleavage and polyadenylation (Birnstiel et al. (1985) Cell 41:349; Proudfoot and Whitelaw (1988) "Termination and 3' end processing of eukaryotic RNA. In Transcription and splicing (ed. B.D. Hames and D.M.

Glover); Proudfoot (1989) Trends Biochem. Sci. 14:105). These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA. Examples of transcription terminator/polyadenylation signals include those derived from SV40 (Sambrook et al (1989) "Expression of cloned genes in cultured mammalian cells." In Molecular Cloning: A Laboratory Manual).

Usually, the above described components, comprising a promoter, polyadenylation signal, and transcription termination sequence are put together into expression constructs. Enhancers, introns with functional splice donor and acceptor sites, and leader sequences may also be included in an expression construct, if desired. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as mammalian cells or bacteria. Mammalian replication systems include those derived from animal viruses, which require trans-acting factors to replicate. For example, plasmids containing the replication systems of papovaviruses, such as SV40 (Gluzman (1981) Cell 23:175) or polyomavirus, replicate to extremely high copy number in the presence of the appropriate viral T antigen. Additional examples of mammalian replicons include those derived from bovine papillomavirus and Epstein-Barr virus. Additionally, the replicon may have two replication systems, thus allowing it to be maintained, for example, in mammalian cells for expression and in a prokaryotic host for cloning and amplification. Examples of such mammalian-bacteria shuttle vectors include pMT2 (Kaufman et al. (1989) Mol. Cell. Biol. 9:946) and pHEBO (Shimizu et al. (1986) Mol. Cell. Biol. 6:1074).

The transformation procedure used depends upon the host to be transformed. Methods for introduction of heterologous polynucleotides into mammalian cells are known in the art and include dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei.

Mammalian cell lines available as hosts for expression are known in the art and include many immortalized cell lines available from the American Type Culture Collection (ATCC), including but not limited to, Chinese hamster ovary (CHO) cells, HeLa cells, baby hamster kidney (BHK) cells, monkey kidney cells (COS), human hepatocellular carcinoma cells (e.g., Hep G2), and a number of other cell lines.

ii. Plant Cellular Expression Systems

There are many plant cell culture and whole plant genetic expression systems known in the art. Exemplary plant cellular genetic expression systems include those described in patents, such as: U.S. 5,693,506; US 5,659,122; and US 5,608,143. Additional examples of genetic expression in plant cell culture has been described by Zenk, Phytochemistry 30:3861-3863 (1991). Descriptions of plant protein signal peptides may be found in addition to the references described above in Vaulcombe et al., Mol. Gen. Genet. 209:33-40 (1987); Chandler et al., Plant Molecular Biology 3:407-418 (1984); Rogers, J. Biol. Chem. 260:3731-3738 (1985); Rothstein et al., Gene 55:353-356 (1987); Whittier et al., Nucleic Acids Research 15:2515-2535 (1987); Wirsel et al., Molecular Microbiology 3:3-14 (1989); Yu et al., Gene 122:247-253 (1992). A description of the regulation of plant gene expression by the phytohormone, gibberellic acid and secreted enzymes induced by gibberellic acid can be found in R.L. Jones and J. MacMillin, Gibberellins: in: Advanced Plant Physiology,. Malcolm B. Wilkins, ed., 1984 Pitman Publishing Limited, London, pp. 21-52. References that describe other metabolically-regulated genes: Sheen, Plant Cell, 2:1027-1038(1990); Maas et al., EMBO J. 9:3447-3452 (1990); Benkel and Hickey, Proc. Natl. Acad. Sci. 84:1337-1339 (1987)

Typically, using techniques known in the art, a desired polynucleotide sequence is inserted into an expression cassette comprising genetic regulatory elements designed for operation in plants. The expression cassette is inserted into a desired expression vector with companion sequences upstream and downstream from the expression cassette suitable for expression in a plant host. The companion sequences will be of plasmid or viral origin and provide necessary characteristics to the vector to permit the vectors to move DNA from an original cloning host, such as bacteria, to the desired plant host. The basic bacterial/plant vector construct will preferably provide a broad host range prokaryote replication origin; a prokaryote selectable marker; and, for Agrobacterium transformations, T DNA sequences for Agrobacterium-mediated transfer to plant chromosomes. Where the heterologous gene is not readily amenable to detection, the construct will preferably also have a selectable marker gene suitable for determining if a plant cell has been transformed. A general review of suitable markers, for example for the members of the grass family, is found in Wilmink and Dons, 1993, *Plant Mol. Biol. Reptr.*, 11(2):165-185.

Sequences suitable for permitting integration of the heterologous sequence into the plant genome are also recommended. These might include transposon sequences and the like for homologous recombination as well as Ti sequences which permit random insertion of a heterologous expression cassette into a plant genome. Suitable prokaryote selectable markers include resistance toward antibiotics such as ampicillin or tetracycline. Other DNA sequences encoding additional functions may also be present in the vector, as is known in the art.

The nucleic acid molecules of the subject invention may be included into an expression cassette for expression of the protein(s) of interest. Usually, there will be only one expression cassette, although two or more are feasible. The recombinant expression cassette will contain in addition to the heterologous protein encoding sequence the following elements, a promoter region, plant 5' untranslated sequences, initiation codon depending upon whether or not the structural gene comes equipped with one, and a transcription and translation termination sequence. Unique restriction enzyme sites at the 5' and 3' ends of the cassette allow for easy insertion into a pre-existing vector.

A heterologous coding sequence may be for any protein relating to the present invention. The sequence encoding the protein of interest will encode a signal peptide which allows processing and translocation of the protein, as appropriate, and will usually lack any sequence which might result in the binding of the desired protein of the invention to a membrane. Since, for the most part, the transcriptional initiation region will be for a gene which is expressed and translocated during germination, by employing the signal peptide which provides for translocation, one may also provide for translocation of the protein of interest. In this way, the protein(s) of interest will be translocated from the cells in which they are expressed and may be efficiently harvested. Typically secretion in seeds are across the aleurone or scutellar epithelium layer into the endosperm of the seed. While it is not required that the protein be secreted from the cells in which the protein is produced, this facilitates the isolation and purification of the recombinant protein.

Since the ultimate expression of the desired gene product will be in a eucaryotic cell it is desirable to determine whether any portion of the cloned gene contains sequences which will be processed out as introns by the host's splicosome machinery. If so, site-directed mutagenesis of the "intron" region may be conducted to prevent losing a portion of the genetic message as a false intron code, Reed and Maniatis, *Cell* 41:95-105, 1985.

The vector can be microinjected directly into plant cells by use of micropipettes to mechanically transfer the recombinant DNA. Crossway, *Mol. Gen. Genet*, 202:179-185, 1985. The genetic material may also be transferred into the plant cell by using polyethylene glycol, Krens, et al., *Nature*, 296, 72-74, 1982. Another method of introduction of nucleic acid segments is high velocity ballistic penetration by small particles with the nucleic acid either within the matrix of small beads or particles, or on the surface, Klein, et al., *Nature*, 327, 70-73, 1987 and Knudsen and Muller, 1991, *Planta*, 185:330-336 teaching particle bombardment of barley endosperm to create transgenic barley. Yet another method of introduction would be fusion of protoplasts with other entities, either minicells, cells, lysosomes or other fusible lipid-surfaced bodies, Fraley, et al., *Proc. Natl. Acad. Sci. USA*, 79, 1859-1863, 1982.

The vector may also be introduced into the plant cells by electroporation. (Fromm et al., *Proc. Natl Acad. Sci. USA* 82:5824, 1985). In this technique, plant protoplasts are electroporated in the presence of plasmids containing the gene construct. Electrical impulses of high field strength reversibly permeabilize biomembranes allowing the introduction of the plasmids. Electroporated plant protoplasts reform the cell wall, divide, and form plant callus.

All plants from which protoplasts can be isolated and cultured to give whole regenerated plants can be transformed by the present invention so that whole plants are recovered which contain the transferred gene. It is known that practically all plants can be regenerated from cultured cells or tissues, including but not limited to all major species of sugarcane, sugar beet, cotton, fruit and other trees, legumes and vegetables. Some suitable plants include, for example, species from the genera Fragaria, Lotus, Medicago, Onobrychis, Trifolium, Trigonella, Vigna, Citrus, Linum, Geranium, Manihot, Daucus, Arabidopsis, Brassica, Raphanus, Sinapis, Atropa, Capsicum, Datura, Hyoscyamus, Lycopersion, Nicotiana, Solanum, Petunia, Digitalis, Majorana, Cichorium, Helianthus, Lactuca, Bromus, Asparagus, Antirrhinum, Hererocallis, Nemesia, Pelargonium, Panicum, Pennisetum, Ranunculus, Senecio, Salpiglossis, Cucumis, Browaalia, Glycine, Lolium, Zea, Triticum, Sorghum, and Datura.

Means for regeneration vary from species to species of plants, but generally a suspension of transformed protoplasts containing copies of the heterologous gene is first provided. Callus tissue is formed and shoots may be induced from callus and subsequently rooted. Alternatively, embryo formation can be induced from the protoplast suspension.

These embryos germinate as natural embryos to form plants. The culture media will generally contain various amino acids and hormones, such as auxin and cytokinins. It is also advantageous to add glutamic acid and proline to the medium, especially for such species as corn and alfalfa. Shoots and roots normally develop simultaneously. Efficient regeneration will depend on the medium, on the genotype, and on the history of the culture. If these three variables are controlled, then regeneration is fully reproducible and repeatable.

In some plant cell culture systems, the desired protein of the invention may be excreted or alternatively, the protein may be extracted from the whole plant. Where the desired protein of the invention is secreted into the medium, it may be collected.

Alternatively, the embryos and embryoless-half seeds or other plant tissue may be mechanically disrupted to release any secreted protein between cells and tissues. The mixture may be suspended in a buffer solution to retrieve soluble proteins. Conventional protein isolation and purification methods will be then used to purify the recombinant protein. Parameters of time, temperature pH, oxygen, and volumes will be adjusted through routine methods to optimize expression and recovery of heterologous protein.

iii. Baculovirus Systems

The polynucleotide encoding the protein can also be inserted into a suitable insect expression vector, and is operably linked to the control elements within that vector. Vector construction employs techniques which are known in the art. Generally, the components of the expression system include a transfer vector, usually a bacterial plasmid, which contains both a fragment of the baculovirus genome, and a convenient restriction site for insertion of the heterologous gene or genes to be expressed; a wild type baculovirus with a sequence homologous to the baculovirus-specific fragment in the transfer vector (this allows for the homologous recombination of the heterologous gene in to the baculovirus genome); and appropriate insect host cells and growth media.

After inserting the DNA sequence encoding the protein into the transfer vector, the vector and the wild type viral genome are transfected into an insect host cell where the vector and viral genome are allowed to recombine. The packaged recombinant virus is expressed and recombinant plaques are identified and purified. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *inter alia*, Invitrogen, San Diego CA ("MaxBac" kit). These techniques are generally known to

those skilled in the art and fully described in Summers and Smith, *Texas Agricultural Experiment Station Bulletin No. 1555* (1987) (hereinafter "Summers and Smith").

Prior to inserting the DNA sequence encoding the protein into the baculovirus genome, the above described components, comprising a promoter, leader (if desired), coding sequence of interest, and transcription termination sequence, are usually assembled into an intermediate transplacement construct (transfer vector). This construct may contain a single gene and operably linked regulatory elements; multiple genes, each with its owned set of operably linked regulatory elements; or multiple genes, regulated by the same set of regulatory elements. Intermediate transplacement constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as a bacterium. The replicon will have a replication system, thus allowing it to be maintained in a suitable host for cloning and amplification.

Currently, the most commonly used transfer vector for introducing foreign genes into AcNPV is pAc373. Many other vectors, known to those of skill in the art, have also been designed. These include, for example, pVL985 (which alters the polyhedrin start codon from ATG to ATT, and which introduces a BamHI cloning site 32 basepairs downstream from the ATT; see Luckow and Summers, *Virology* (1989) 17:31.

The plasmid usually also contains the polyhedrin polyadenylation signal (Miller et al. (1988) Ann. Rev. Microbiol., 42:177) and a prokaryotic ampicillin-resistance (amp) gene and origin of replication for selection and propagation in E. coli.

Baculovirus transfer vectors usually contain a baculovirus promoter. A baculovirus promoter is any DNA sequence capable of binding a baculovirus RNA polymerase and initiating the downstream (5' to 3') transcription of a coding sequence (e.g., structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site and a transcription initiation site. A baculovirus transfer vector may also have a second domain called an enhancer, which, if present, is usually distal to the structural gene. Expression may be either regulated or constitutive.

Structural genes, abundantly transcribed at late times in a viral infection cycle, provide particularly useful promoter sequences. Examples include sequences derived from the gene encoding the viral polyhedron protein, Friesen et al., (1986) "The Regulation of Baculovirus Gene Expression," in: *The Molecular Biology of Baculoviruses* (ed. Walter Doerfler); EPO

Publ. Nos. 127 839 and 155 476; and the gene encoding the p10 protein, Vlak et al., (1988), J. Gen. Virol. 69:765.

DNA encoding suitable signal sequences can be derived from genes for secreted insect or baculovirus proteins, such as the baculovirus polyhedrin gene (Carbonell et al. (1988) *Gene*, 73:409). Alternatively, since the signals for mammalian cell posttranslational modifications (such as signal peptide cleavage, proteolytic cleavage, and phosphorylation) appear to be recognized by insect cells, and the signals required for secretion and nuclear accumulation also appear to be conserved between the invertebrate cells and vertebrate cells, leaders of non-insect origin, such as those derived from genes encoding human (alpha) α-interferon, Maeda et al., (1985), *Nature 315*:592; human gastrin-releasing peptide, Lebacq-Verheyden et al., (1988), *Molec. Cell. Biol. 8*:3129; human IL-2, Smith et al., (1985) *Proc. Nat'l Acad. Sci. USA*, 82:8404; mouse IL-3, (Miyajima et al., (1987) *Gene 58*:273; and human glucocerebrosidase, Martin et al. (1988) *DNA*, 7:99, can also be used to provide for secretion in insects.

A recombinant polypeptide or polyprotein may be expressed intracellularly or, if it is expressed with the proper regulatory sequences, it can be secreted. Good intracellular expression of nonfused foreign proteins usually requires heterologous genes that ideally have a short leader sequence containing suitable translation initiation signals preceding an ATG start signal. If desired, methionine at the N-terminus may be cleaved from the mature proteins by *in vitro* incubation with cyanogen bromide.

Alternatively, recombinant polyproteins or proteins which are not naturally secreted can be secreted from the insect cell by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provides for secretion of the foreign protein in insects. The leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the translocation of the protein into the endoplasmic reticulum.

After insertion of the DNA sequence and/or the gene encoding the expression product precursor of the protein, an insect cell host is co-transformed with the heterologous DNA of the transfer vector and the genomic DNA of wild type baculovirus -- usually by co-transfection. The promoter and transcription termination sequence of the construct will usually comprise a 2-5kb section of the baculovirus genome. Methods for introducing heterologous DNA into the desired site in the baculovirus virus are known in the art. (See

Summers and Smith supra; Ju et al. (1987); Smith et al., Mol. Cell. Biol. (1983) 3:2156; and Luckow and Summers (1989)). For example, the insertion can be into a gene such as the polyhedrin gene, by homologous double crossover recombination; insertion can also be into a restriction enzyme site engineered into the desired baculovirus gene. Miller et al., (1989), Bioessays 4:91. The DNA sequence, when cloned in place of the polyhedrin gene in the expression vector, is flanked both 5' and 3' by polyhedrin-specific sequences and is positioned downstream of the polyhedrin promoter.

The newly formed baculovirus expression vector is subsequently packaged into an infectious recombinant baculovirus. Homologous recombination occurs at low frequency (between about 1% and about 5%); thus, the majority of the virus produced after cotransfection is still wild-type virus. Therefore, a method is necessary to identify recombinant viruses. An advantage of the expression system is a visual screen allowing recombinant viruses to be distinguished. The polyhedrin protein, which is produced by the native virus, is produced at very high levels in the nuclei of infected cells at late times after viral infection. Accumulated polyhedrin protein forms occlusion bodies that also contain embedded particles. These occlusion bodies, up to 15 μm in size, are highly refractile, giving them a bright shiny appearance that is readily visualized under the light microscope. Cells infected with recombinant viruses lack occlusion bodies. To distinguish recombinant virus from wild-type virus, the transfection supernatant is plaqued onto a monolayer of insect cells by techniques known to those skilled in the art. Namely, the plaques are screened under the light microscope for the presence (indicative of wild-type virus) or absence (indicative of recombinant virus) of occlusion bodies. Current Protocols in Microbiology Vol. 2 (Ausubel et al. eds) at 16.8 (Supp. 10, 1990); Summers and Smith, supra; Miller et al. (1989).

Recombinant baculovirus expression vectors have been developed for infection into several insect cells. For example, recombinant baculoviruses have been developed for, inter alia: Aedes aegypti, Autographa californica, Bombyx mori, Drosophila melanogaster, Spodoptera frugiperda, and Trichoplusia ni (PCT Pub. No. WO 89/046699; Carbonell et al., (1985) J. Virol. 56:153; Wright (1986) Nature 321:718; Smith et al., (1983) Mol. Cell. Biol. 3:2156; and see generally, Fraser, et al. (1989) In Vitro Cell. Dev. Biol. 25:225).

Cells and cell culture media are commercially available for both direct and fusion expression of heterologous polypeptides in a baculovirus/expression system; cell culture technology is generally known to those skilled in the art. See, e.g., Summers and Smith supra.

The modified insect cells may then be grown in an appropriate nutrient medium, which allows for stable maintenance of the plasmid(s) present in the modified insect host. Where the expression product gene is under inducible control, the host may be grown to high density, and expression induced. Alternatively, where expression is constitutive, the product will be continuously expressed into the medium and the nutrient medium must be continuously circulated, while removing the product of interest and augmenting depleted nutrients. The product may be purified by such techniques as chromatography, e.g., HPLC, affinity chromatography, ion exchange chromatography, etc.; electrophoresis; density gradient centrifugation; solvent extraction, or the like. As appropriate, the product may be further purified, as required, so as to remove substantially any insect proteins which are also secreted in the medium or result from lysis of insect cells, so as to provide a product which is at least substantially free of host debris, e.g., proteins, lipids and polysaccharides.

In order to obtain protein expression, recombinant host cells derived from the transformants are incubated under conditions which allow expression of the recombinant protein encoding sequence. These conditions will vary, dependent upon the host cell selected. However, the conditions are readily ascertainable to those of ordinary skill in the art, based upon what is known in the art.

iv. Bacterial Systems

Bacterial expression techniques are known in the art. A bacterial promoter is any DNA sequence capable of binding bacterial RNA polymerase and initiating the downstream (3') transcription of a coding sequence (e.g. structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site and a transcription initiation site. A bacterial promoter may also have a second domain called an operator, that may overlap an adjacent RNA polymerase binding site at which RNA synthesis begins. The operator permits negative regulated (inducible) transcription, as a gene repressor protein may bind the operator and thereby inhibit transcription of a specific gene. Constitutive expression may occur in the absence of negative regulatory elements, such as the operator. In addition, positive regulation may be achieved by a gene activator protein binding sequence, which, if present is usually proximal (5') to the RNA polymerase binding sequence. An example of a gene activator protein is the catabolite activator protein (CAP), which helps

initiate transcription of the lac operon in Escherichia coli (E. coli) (Raibaud et al. (1984) Annu. Rev. Genet. 18:173). Regulated expression may therefore be either positive or negative, thereby either enhancing or reducing transcription.

Sequences encoding metabolic pathway enzymes provide particularly useful promoter sequences. Examples include promoter sequences derived from sugar metabolizing enzymes, such as galactose, lactose (lac) (Chang et al. (1977) Nature 198:1056), and maltose. Additional examples include promoter sequences derived from biosynthetic enzymes such as tryptophan (trp) (Goeddel et al. (1980) Nuc. Acids Res. 8:4057; Yelverton et al. (1981) Nucl. Acids Res. 9:731; U.S. Patent 4,738,921; EPO Publ. Nos. 036 776 and 121 775). The betalactamase (bla) promoter system (Weissmann (1981) "The cloning of interferon and other mistakes." In Interferon 3 (ed. I. Gresser)), bacteriophage lambda PL (Shimatake et al. (1981) Nature 292:128) and T5 (U.S. Patent 4,689,406) promoter systems also provide useful promoter sequences.

In addition, synthetic promoters which do not occur in nature also function as bacterial promoters. For example, transcription activation sequences of one bacterial or bacteriophage promoter may be joined with the operon sequences of another bacterial or bacteriophage promoter, creating a synthetic hybrid promoter (U.S. Patent 4,551,433). For example, the *tac* promoter is a hybrid *trp-lac* promoter comprised of both *trp* promoter and *lac* operon sequences that is regulated by the *lac* repressor (Amann *et al.* (1983) *Gene 25*:167; de Boer *et al.* (1983) *Proc. Natl. Acad. Sci. 80*:21). Furthermore, a bacterial promoter can include naturally occurring promoters of non-bacterial origin that have the ability to bind bacterial RNA polymerase and initiate transcription. A naturally occurring promoter of non-bacterial origin can also be coupled with a compatible RNA polymerase to produce high levels of expression of some genes in prokaryotes. The bacteriophage T7 RNA polymerase/promoter system is an example of a coupled promoter system (Studier *et al.* (1986) *J. Mol. Biol. 189*:113; Tabor *et al.* (1985) *Proc Natl. Acad. Sci. 82*:1074). In addition, a hybrid promoter can also be comprised of a bacteriophage promoter and an *E. coli* operator region (EPO Publ. No. 267 851).

In addition to a functioning promoter sequence, an efficient ribosome binding site is also useful for the expression of foreign genes in prokaryotes. In *E. coli*, the ribosome binding site is called the Shine-Dalgarno (SD) sequence and includes an initiation codon (ATG) and a sequence 3-9 nucleotides in length located 3-11 nucleotides upstream of the initiation codon

(Shine et al. (1975) Nature 254:34). The SD sequence is thought to promote binding of mRNA to the ribosome by the pairing of bases between the SD sequence and the 3' end of E. coli 16S rRNA (Steitz et al. (1979) "Genetic signals and nucleotide sequences in messenger RNA." In Biological Regulation and Development: Gene Expression (ed. R.F. Goldberger)). To express eukaryotic genes and prokaryotic genes with weak ribosome-binding site, it is often necessary to optimize the distance between the SD sequence and the ATG of the eukaryotic gene (Sambrook et al. (1989) "Expression of cloned genes in Escherichia coli." In Molecular Cloning: A Laboratory Manual).

A DNA molecule may be expressed intracellularly. A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus will always be a methionine, which is encoded by the ATG start codon. If desired, methionine at the N-terminus may be cleaved from the protein by *in vitro* incubation with cyanogen bromide or by either *in vivo* or *in vitro* incubation with a bacterial methionine N-terminal peptidase (EPO Publ. No. 219 237).

Fusion proteins provide an alternative to direct expression. Usually, a DNA sequence encoding the N-terminal portion of an endogenous bacterial protein, or other stable protein, is fused to the 5' end of heterologous coding sequences. Upon expression, this construct will provide a fusion of the two amino acid sequences. For example, the bacteriophage lambda cell gene can be linked at the 5' terminus of a foreign gene and expressed in bacteria. The resulting fusion protein preferably retains a site for a processing enzyme (factor Xa) to cleave the bacteriophage protein from the foreign gene (Nagai et al. (1984) Nature 309:810). Fusion proteins can also be made with sequences from the lacZ (Jia et al. (1987) Gene 60:197), trpE (Allen et al. (1987) J. Biotechnol. 5:93; Makoff et al. (1989) J. Gen. Microbiol. 135:11), and Chey (EPO Publ. No. 324 647) genes. The DNA sequence at the junction of the two amino acid sequences may or may not encode a cleavable site. Another example is a ubiquitin fusion protein. Such a fusion protein is made with the ubiquitin region that preferably retains a site for a processing enzyme (e.g. ubiquitin specific processing-protease) to cleave the ubiquitin from the foreign protein. Through this method, native foreign protein can be isolated (Miller et al. (1989) Bio/Technology 7:698).

Alternatively, foreign proteins can also be secreted from the cell by creating chimeric DNA molecules that encode a fusion protein comprised of a signal peptide sequence fragment that provides for secretion of the foreign protein in bacteria (U.S. Patent 4,336,336). The

signal sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell. The protein is either secreted into the growth media (gram-positive bacteria) or into the periplasmic space, located between the inner and outer membrane of the cell (gram-negative bacteria). Preferably there are processing sites, which can be cleaved either *in vivo* or *in vitro* encoded between the signal peptide fragment and the foreign gene.

DNA encoding suitable signal sequences can be derived from genes for secreted bacterial proteins, such as the *E. coli* outer membrane protein gene (*ompA*) (Masui *et al.* (1983), in: *Experimental Manipulation of Gene Expression*; Ghrayeb *et al.* (1984) *EMBO J.* 3:2437) and the *E. coli* alkaline phosphatase signal sequence (*phoA*) (Oka *et al.* (1985) *Proc. Natl. Acad. Sci. 82*:7212). As an additional example, the signal sequence of the alpha-amylase gene from various Bacillus strains can be used to secrete heterologous proteins from *B. subtilis* (Palva *et al.* (1982) *Proc. Natl. Acad. Sci. USA 79*:5582; EPO Publ. No. 244 042).

Usually, transcription termination sequences recognized by bacteria are regulatory regions located 3' to the translation stop codon, and thus together with the promoter flank the coding sequence. These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA. Transcription termination sequences frequently include DNA sequences of about 50 nucleotides capable of forming stem loop structures that aid in terminating transcription. Examples include transcription termination sequences derived from genes with strong promoters, such as the *trp* gene in *E. coli* as well as other biosynthetic genes.

Usually, the above described components, comprising a promoter, signal sequence (if desired), coding sequence of interest, and transcription termination sequence, are put together into expression constructs. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as bacteria. The replicon will have a replication system, thus allowing it to be maintained in a prokaryotic host either for expression or for cloning and amplification. In addition, a replicon may be either a high or low copy number plasmid. A high copy number plasmid will generally have a copy number ranging from about 5 to about 200, and usually about 10 to about 150. A host containing a high copy number plasmid will preferably contain at least about 10, and more preferably at least about 20 plasmids. Either a high or low copy number

vector may be selected, depending upon the effect of the vector and the foreign protein on the host.

Alternatively, the expression constructs can be integrated into the bacterial genome with an integrating vector. Integrating vectors usually contain at least one sequence homologous to the bacterial chromosome that allows the vector to integrate. Integrations appear to result from recombinations between homologous DNA in the vector and the bacterial chromosome. For example, integrating vectors constructed with DNA from various Bacillus strains integrate into the Bacillus chromosome (EPO Publ. No. 127 328). Integrating vectors may also be comprised of bacteriophage or transposon sequences.

Usually, extrachromosomal and integrating expression constructs may contain selectable markers to allow for the selection of bacterial strains that have been transformed. Selectable markers can be expressed in the bacterial host and may include genes which render bacteria resistant to drugs such as ampicillin, chloramphenicol, erythromycin, kanamycin (neomycin), and tetracycline (Davies et al. (1978) Annu. Rev. Microbiol. 32:469). Selectable markers may also include biosynthetic genes, such as those in the histidine, tryptophan, and leucine biosynthetic pathways.

Alternatively, some of the above described components can be put together in transformation vectors. Transformation vectors are usually comprised of a selectable market that is either maintained in a replicon or developed into an integrating vector, as described above.

Expression and transformation vectors, either extra-chromosomal replicons or integrating vectors, have been developed for transformation into many bacteria. For example, expression vectors have been developed for, *inter alia*, the following bacteria: Bacillus subtilis (Palva *et al.* (1982) *Proc. Natl. Acad. Sci. USA 79*:5582; EPO Publ. Nos. 036 259 and 063 953; PCT Publ. No. WO 84/04541), Escherichia coli (Shimatake *et al.* (1981) *Nature 292*:128; Amann *et al.* (1985) *Gene 40*:183; Studier *et al.* (1986) *J. Mol. Biol. 189*:113; EPO Publ. Nos. 036 776, 136 829 and 136 907), Streptococcus cremoris (Powell *et al.* (1988) *Appl. Environ. Microbiol. 54*:655); Streptococcus lividans (Powell *et al.* (1988) *Appl. Environ. Microbiol. 54*:655), Streptomyces lividans (U.S. Patent 4,745,056).

Methods of introducing exogenous DNA into bacterial hosts are well-known in the art, and usually include either the transformation of bacteria treated with CaCl₂ or other agents, such as divalent cations and DMSO. DNA can also be introduced into bacterial cells by

electroporation. Transformation procedures usually vary with the bacterial species to be transformed. (See e.g., use of Bacillus: Masson et al. (1989) FEMS Microbiol. Lett. 60:273; Palva et al. (1982) Proc. Natl. Acad. Sci. USA 79:5582; EPO Publ. Nos. 036 259 and 063 953; PCT Publ. No. WO 84/04541; use of Campylobacter: Miller et al. (1988) Proc. Natl. Acad. Sci. 85:856; and Wang et al. (1990) J. Bacteriol. 172:949; use of Escherichia coli: Cohen et al. (1973) Proc. Natl. Acad. Sci. 69:2110; Dower et al. (1988) Nucleic Acids Res. 16:6127; Kushner (1978) "An improved method for transformation of Escherichia coli with ColE1derived plasmids. In Genetic Engineering: Proceedings of the International Symposium on Genetic Engineering (eds. H.W. Boyer and S. Nicosia); Mandel et al. (1970) J. Mol. Biol. 53:159; Taketo (1988) Biochim. Biophys. Acta 949:318; use of Lactobacillus: Chassy et al. (1987) FEMS Microbiol. Lett. 44:173; use of Pseudomonas: Fiedler et al. (1988) Anal. Biochem 170:38; use of Staphylococcus: Augustin et al. (1990) FEMS Microbiol. Lett. 66:203; use of Streptococcus: Barany et al. (1980) J. Bacteriol. 144:698; Harlander (1987) "Transformation of Streptococcus lactis by electroporation, in: Streptococcal Genetics (ed. J. Ferretti and R. Curtiss III); Perry et al. (1981) Infect. Immun. 32:1295; Powell et al. (1988) Appl. Environ. Microbiol. 54:655; Somkuti et al. (1987) Proc. 4th Evr. Cong. Biotechnology *I*:412.

v. Yeast Expression

Yeast expression systems are also known to one of ordinary skill in the art. A yeast promoter is any DNA sequence capable of binding yeast RNA polymerase and initiating the downstream (3') transcription of a coding sequence (e.g. structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site (the "TATA Box") and a transcription initiation site. A yeast promoter may also have a second domain called an upstream activator sequence (UAS), which, if present, is usually distal to the structural gene. The UAS permits regulated (inducible) expression. Constitutive expression occurs in the absence of a UAS. Regulated expression may be either positive or negative, thereby either enhancing or reducing transcription.

Yeast is a fermenting organism with an active metabolic pathway, therefore sequences encoding enzymes in the metabolic pathway provide particularly useful promoter sequences.

Examples include alcohol dehydrogenase (ADH) (EPO Publ. No. 284 044), enolase, glucokinase, glucose-6-phosphate isomerase, glyceraldehyde-3-phosphate-dehydrogenase (GAP or GAPDH), hexokinase, phosphofructokinase, 3-phosphoglycerate mutase, and pyruvate kinase (PyK) (EPO Publ. No. 329 203). The yeast *PHO5* gene, encoding acid phosphatase, also provides useful promoter sequences (Myanohara *et al.* (1983) *Proc. Natl. Acad. Sci. USA 80*:1).

In addition, synthetic promoters which do not occur in nature also function as yeast promoters. For example, UAS sequences of one yeast promoter may be joined with the transcription activation region of another yeast promoter, creating a synthetic hybrid promoter. Examples of such hybrid promoters include the ADH regulatory sequence linked to the GAP transcription activation region (U.S. Patent Nos. 4,876,197 and 4,880,734). Other examples of hybrid promoters include promoters which consist of the regulatory sequences of either the ADH2, GAL4, GAL10, OR PHO5 genes, combined with the transcriptional activation region of a glycolytic enzyme gene such as GAP or PyK (EPO Publ. No. 164 556). Furthermore, a yeast promoter can include naturally occurring promoters of non-yeast origin that have the ability to bind yeast RNA polymerase and initiate transcription. Examples of such promoters include, inter alia, (Cohen et al. (1980) Proc. Natl. Acad. Sci. USA 77:1078; Henikoff et al. (1981) Nature 283:835; Hollenberg et al. (1981) Curr. Topics Microbiol. Immunol. 96:119; Hollenberg et al. (1979) "The Expression of Bacterial Antibiotic Resistance Genes in the Yeast Saccharomyces cerevisiae," in: Plasmids of Medical, Environmental and Commercial Importance (eds. K.N. Timmis and A. Puhler); Mercerau-Puigalon et al. (1980) Gene 11:163; Panthier et al. (1980) Curr. Genet. 2:109;).

A DNA molecule may be expressed intracellularly in yeast. A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus of the recombinant protein will always be a methionine, which is encoded by the ATG start codon. If desired, methionine at the N-terminus may be cleaved from the protein by in vitro incubation with cyanogen bromide.

Fusion proteins provide an alternative for yeast expression systems, as well as in mammalian, plant, baculovirus, and bacterial expression systems. Usually, a DNA sequence encoding the N-terminal portion of an endogenous yeast protein, or other stable protein, is fused to the 5' end of heterologous coding sequences. Upon expression, this construct will provide a fusion of the two amino acid sequences. For example, the yeast or human

superoxide dismutase (SOD) gene, can be linked at the 5' terminus of a foreign gene and expressed in yeast. The DNA sequence at the junction of the two amino acid sequences may or may not encode a cleavable site. See e.g., EPO Publ. No. 196056. Another example is a ubiquitin fusion protein. Such a fusion protein is made with the ubiquitin region that preferably retains a site for a processing enzyme (e.g. ubiquitin-specific processing protease) to cleave the ubiquitin from the foreign protein. Through this method, therefore, native foreign protein can be isolated (e.g., WO88/024066).

Alternatively, foreign proteins can also be secreted from the cell into the growth media by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provide for secretion in yeast of the foreign protein. Preferably, there are processing sites encoded between the leader fragment and the foreign gene that can be cleaved either *in vivo* or *in vitro*. The leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell.

DNA encoding suitable signal sequences can be derived from genes for secreted yeast proteins, such as the yeast invertase gene (EPO Publ. No. 012 873; JPO Publ. No. 62:096,086) and the A-factor gene (U.S. Patent 4,588,684). Alternatively, leaders of non-yeast origin, such as an interferon leader, exist that also provide for secretion in yeast (EPO Publ. No. 060 057).

A preferred class of secretion leaders are those that employ a fragment of the yeast alpha-factor gene, which contains both a "pre" signal sequence, and a "pro" region. The types of alpha-factor fragments that can be employed include the full-length pre-pro alpha factor leader (about 83 amino acid residues) as well as truncated alpha-factor leaders (usually about 25 to about 50 amino acid residues) (U.S. Patent Nos. 4,546,083 and 4,870,008; EPO Publ. No. 324 274). Additional leaders employing an alpha-factor leader fragment that provides for secretion include hybrid alpha-factor leaders made with a presequence of a first yeast, but a pro-region from a second yeast alphafactor. (See e.g., PCT Publ. No. WO 89/02463.)

Usually, transcription termination sequences recognized by yeast are regulatory regions located 3' to the translation stop codon, and thus together with the promoter flank the coding sequence. These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA. Examples of transcription terminator

sequence and other yeast-recognized termination sequences, such as those coding for glycolytic enzymes.

Usually, the above described components, comprising a promoter, leader (if desired), coding sequence of interest, and transcription termination sequence, are put together into expression constructs. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as yeast or bacteria. The replicon may have two replication systems, thus allowing it to be maintained, for example, in yeast for expression and in a prokaryotic host for cloning and amplification. Examples of such yeast-bacteria shuttle vectors include YEp24 (Botstein et al. (1979) Gene 8:17-24), pCl/1 (Brake et al. (1984) Proc. Natl. Acad. Sci USA 81:4642-4646), and YRp17 (Stinchcomb et al. (1982) J. Mol. Biol. 158:157). In addition, a replicon may be either a high or low copy number plasmid. A high copy number plasmid will generally have a copy number ranging from about 5 to about 200, and usually about 10 to about 150. A host containing a high copy number plasmid will preferably have at least about 10, and more preferably at least about 20. Enter a high or low copy number vector may be selected, depending upon the effect of the vector and the foreign protein on the host. See e.g., Brake et al., supra.

Alternatively, the expression constructs can be integrated into the yeast genome with an integrating vector. Integrating vectors usually contain at least one sequence homologous to a yeast chromosome that allows the vector to integrate, and preferably contain two homologous sequences flanking the expression construct. Integrations appear to result from recombinations between homologous DNA in the vector and the yeast chromosome (Orr-Weaver et al. (1983) Methods in Enzymol. 101:228-245). An integrating vector may be directed to a specific locus in yeast by selecting the appropriate homologous sequence for inclusion in the vector. See Orr-Weaver et al., supra. One or more expression construct may integrate, possibly affecting levels of recombinant protein produced (Rine et al. (1983) Proc. Natl. Acad. Sci. USA 80:6750). The chromosomal sequences included in the vector can occur either as a single segment in the vector, which results in the integration of the entire vector, or two segments homologous to adjacent segments in the chromosome and flanking the expression construct in the vector, which can result in the stable integration of only the expression construct.

Usually, extrachromosomal and integrating expression constructs may contain selectable markers to allow for the selection of yeast strains that have been transformed. Selectable markers may include biosynthetic genes that can be expressed in the yeast host, such as ADE2, HIS4, LEU2, TRP1, and ALG7, and the G418 resistance gene, which confer resistance in yeast cells to tunicamycin and G418, respectively. In addition, a suitable selectable marker may also provide yeast with the ability to grow in the presence of toxic compounds, such as metal. For example, the presence of CUP1 allows yeast to grow in the presence of copper ions (Butt et al. (1987) Microbiol, Rev. 51:351).

Alternatively, some of the above described components can be put together into transformation vectors. Transformation vectors are usually comprised of a selectable marker that is either maintained in a replicon or developed into an integrating vector, as described above.

Expression and transformation vectors, either extrachromosomal replicons or integrating vectors, have been developed for transformation into many yeasts. For example, expression vectors and methods of introducing exogenous DNA into yeast hosts have been developed for, inter alia, the following yeasts: Candida albicans (Kurtz, et al. (1986) Mol. Cell. Biol. 6:142); Candida maltosa (Kunze, et al. (1985) J. Basic Microbiol. 25:141); Hansenula polymorpha (Gleeson, et al. (1986) J. Gen. Microbiol. 132:3459; Roggenkamp et al. (1986) Mol. Gen. Genet. 202:302); Kluyveromyces fragilis (Das, et al. (1984) J. Bacteriol. 158:1165); Kluyveromyces lactis (De Louvencourt et al. (1983) J. Bacteriol. 154:737; Van den Berg et al. (1990) Bio/Technology 8:135); Pichia guillerimondii (Kunze et al. (1985) J. Basic Microbiol. 25:141); Pichia pastoris (Cregg, et al. (1985) Mol. Cell. Biol. 5:3376; U.S. Patent Nos. 4,837,148 and 4,929,555); Saccharomyces cerevisiae (Hinnen et al. (1978) Proc. Natl. Acad. Sci. USA 75:1929; Ito et al. (1983) J. Bacteriol. 153:163); Schizosaccharomyces pombe (Beach and Nurse (1981) Nature 300:706); and Yarrowia lipolytica (Davidow, et al. (1985) Curr. Genet. 10:380471 Gaillardin, et al. (1985) Curr. Genet. 10:49).

Methods of introducing exogenous DNA into yeast hosts are well-known in the art, and usually include either the transformation of spheroplasts or of intact yeast cells treated with alkali cations. Transformation procedures usually vary with the yeast species to be transformed. See e.g., [Kurtz et al. (1986) Mol. Cell. Biol. 6:142; Kunze et al. (1985) J. Basic Microbiol. 25:141; Candida]; [Gleeson et al. (1986) J. Gen. Microbiol. 132:3459; Roggenkamp et al. (1986) Mol. Gen. Genet. 202:302; Hansenula]; [Das et al. (1984) J.

Bacteriol. 158:1165; De Louvencourt et al. (1983) J. Bacteriol. 154:1165; Van den Berg et al. (1990) Bio/Technology 8:135; Kluyveromyces]; [Cregg et al. (1985) Mol. Cell. Biol. 5:3376; Kunze et al. (1985) J. Basic Microbiol. 25:141; U.S. Patent Nos. 4,837,148 and 4,929,555; Pichia]; [Hinnen et al. (1978) Proc. Natl. Acad. Sci. USA 75;1929; Ito et al. (1983) J. Bacteriol. 153:163 Saccharomyces]; [Beach and Nurse (1981) Nature 300:706; Schizosaccharomyces]; [Davidow et al. (1985) Curr. Genet. 10:39; Gaillardin et al. (1985) Curr. Genet. 10:49; Yarrowia].

Definitions

A composition containing X is "substantially free of" Y when at least 85% by weight of the total X+Y in the composition is X. Preferably, X comprises at least about 90% by weight of the total of X+Y in the composition, more preferably at least about 95% or even 99% by weight.

A "conserved" Neisseria amino acid fragment or protein is one that is present in a particular Neisserial protein in at least x% of Neisseria. The value of x may be 50% or more, e.g., 66%, 75%, 80%, 90%, 95% or even 100% (i.e. the amino acid is found in the protein in question in all Neisseria). In order to determine whether an animo acid is "conserved" in a particular Neisserial protein, it is necessary to compare that amino acid residue in the sequences of the protein in question from a plurality of different Neisseria (a reference population). The reference population may include a number of different Neisseria species or may include a single species. The reference population may include a number of different serogroups of a particular species or a single serogroup. A preferred reference population consists of the 5 most common Neisseria strains.

The term "heterologous" refers to two biological components that are not found together in nature. The components may be host cells, genes, or regulatory regions, such as promoters. Although the heterologous components are not found together in nature, they can function together, as when a promoter heterologous to a gene is operably linked to the gene. Another example is where a Neisserial sequence is heterologous to a mouse host cell.

"Epitope" means antigenic determinant, and may elicit a cellular and/or humoral response.

Conditions for "high stringency" are 65 degrees C in 0.1 xSSC 0.5% SDS solution.

An "origin of replication" is a polynucleotide sequence that initiates and regulates replication of polynucleotides, such as an expression vector. The origin of replication behaves as an autonomous unit of polynucleotide replication within a cell, capable of replication under its own control. An origin of replication may be needed for a vector to replicate in a particular host cell. With certain origins of replication, an expression vector can be reproduced at a high copy number in the presence of the appropriate proteins within the cell. Examples of origins are the autonomously replicating sequences, which are effective in yeast; and the viral T-antigen, effective in COS-7 cells.

A "mutant" sequence is defined as a DNA, RNA or amino acid sequence differing from but having homology with the native or disclosed sequence. Depending on the particular sequence, the degree of homology (sequence identity) between the native or disclosed sequence and the mutant sequence is preferably greater than 50% (e.g., 60%, 70%, 80%, 90%, 95%, 99% or more) which is calculated as described above. As used herein, an "allelic variant" of a nucleic acid molecule, or region, for which nucleic acid sequence is provided herein is a nucleic acid molecule, or region, that occurs at essentially the same locus in the genome of another or second isolate, and that, due to natural variation caused by, for example, mutation or recombination, has a similar but not identical nucleic acid sequence. A coding region allelic variant typically encodes a protein having similar activity to that of the protein encoded by the gene to which it is being compared. An allelic variant can also comprise an alteration in the 5' or 3' untranslated regions of the gene, such as in regulatory control regions. (see, for example, U.S. Patent 5,753,235).

Antibodies

As used herein, the term "antibody" refers to a polypeptide or group of polypeptides composed of at least one antibody combining site. An "antibody combining site" is the three-dimensional binding space with an internal surface shape and charge distribution complementary to the features of an epitope of an antigen, which allows a binding of the antibody with the antigen. "Antibody" includes, for example, vertebrate antibodies, hybrid antibodies, chimeric antibodies, humanized antibodies, altered antibodies, univalent antibodies, Fab proteins, and single domain antibodies.

Antibodies against the proteins of the invention are useful for affinity chromatography, immunoassays, and distinguishing/identifying Neisseria menB proteins.

Antibodies elicited against the proteins of the present invention bind to antigenic polypeptides or proteins or protein fragments that are present and specifically associated with strains of *Neisseria meningitidis* menB. In some instances, these antigens may be associated with specific strains, such as those antigens specific for the menB strains. The antibodies of the invention may be immobilized to a matrix and utilized in an immunoassay or on an affinity chromatography column, to enable the detection and/or separation of polypeptides, proteins or protein fragments or cells comprising such polypeptides, proteins or protein fragments.

Alternatively, such polypeptides, proteins or protein fragments may be immobilized so as to detect antibodies bindably specific thereto.

Antibodies to the proteins of the invention, both polyclonal and monoclonal, may be prepared by conventional methods. In general, the protein is first used to immunize a suitable animal, preferably a mouse, rat, rabbit or goat. Rabbits and goats are preferred for the preparation of polyclonal sera due to the volume of serum obtainable, and the availability of labeled anti-rabbit and anti-goat antibodies. Immunization is generally performed by mixing or emulsifying the protein in saline, preferably in an adjuvant such as Freund's complete adjuvant, and injecting the mixture or emulsion parenterally (generally subcutaneously or intramuscularly). A dose of 50-200 µg/injection is typically sufficient. Immunization is generally boosted 2-6 weeks later with one or more injections of the protein in saline, preferably using Freund's incomplete adjuvant. One may alternatively generate antibodies by in vitro immunization using methods known in the art, which for the purposes of this invention is considered equivalent to in vivo immunization. Polyclonal antisera is obtained by bleeding the immunized animal into a glass or plastic container, incubating the blood at 25°C for one hour, followed by incubating at 4°C for 2-18 hours. The serum is recovered by centrifugation (e.g., 1,000g for 10 minutes). About 20-50 ml per bleed may be obtained from rabbits.

Monoclonal antibodies are prepared using the standard method of Kohler & Milstein (Nature (1975) 256:495-96), or a modification thereof. Typically, a mouse or rat is immunized as described above. However, rather than bleeding the animal to extract serum, the spleen (and optionally several large lymph nodes) is removed and dissociated into single cells. If desired, the spleen cells may be screened (after removal of nonspecifically adherent cells) by applying a cell suspension to a plate or well coated with the protein antigen. B-cells that express membrane-bound immunoglobulin specific for the antigen bind to the plate, and

are not rinsed away with the rest of the suspension. Resulting B-cells, or all dissociated spleen cells, are then induced to fuse with myeloma cells to form hybridomas, and are cultured in a selective medium (e.g., hypoxanthine, aminopterin, thymidine medium, "HAT"). The resulting hybridomas are plated by limiting dilution, and are assayed for the production of antibodies which bind specifically to the immunizing antigen (and which do not bind to unrelated antigens). The selected MAb-secreting hybridomas are then cultured either *in vitro* (e.g., in tissue culture bottles or hollow fiber reactors), or *in vivo* (as ascites in mice).

If desired, the antibodies (whether polyclonal or monoclonal) may be labeled using conventional techniques. Suitable labels include fluorophores, chromophores, radioactive atoms (particularly ³²P and ¹²⁵I), electron-dense reagents, enzymes, and ligands having specific binding partners. Enzymes are typically detected by their activity. For example, horseradish peroxidase is usually detected by its ability to convert 3,3',5,5'-tetramethylbenzidine (TMB) to a blue pigment, quantifiable with a spectrophotometer. "Specific binding partner" refers to a protein capable of binding a ligand molecule with high specificity, as for example in the case of an antigen and a monoclonal antibody specific therefor. Other specific binding partners include biotin and avidin or streptavidin, IgG and protein A, and the numerous receptor-ligand couples known in the art. It should be understood that the above description is not meant to categorize the various labels into distinct classes, as the same label may serve in several different modes. For example, 125I may serve as a radioactive label or as an electron-dense reagent. HRP may serve as enzyme or as antigen for a MAb. Further, one may combine various labels for desired effect. For example, MAbs and avidin also require labels in the practice of this invention: thus, one might label a MAb with biotin, and detect its presence with avidin labeled with 125 I, or with an anti-biotin MAb labeled with HRP. Other permutations and possibilities will be readily apparent to those of ordinary skill in the art, and are considered as equivalents within the scope of the instant invention.

Antigens, immunogens, polypeptides, proteins or protein fragments of the present invention elicit formation of specific binding partner antibodies. These antigens, immunogens, polypeptides, proteins or protein fragments of the present invention comprise immunogenic compositions of the present invention. Such immunogenic compositions may further comprise or include adjuvants, carriers, or other compositions that promote or enhance

or stabilize the antigens, polypeptides, proteins or protein fragments of the present invention. Such adjuvants and carriers will be readily apparent to those of ordinary skill in the art.

Pharmaceutical Compositions

Pharmaceutical compositions can comprise (include) either polypeptides, antibodies, or nucleic acid of the invention. The pharmaceutical compositions will comprise a therapeutically effective amount of either polypeptides, antibodies, or polynucleotides of the claimed invention.

The term "therapeutically effective amount" as used herein refers to an amount of a therapeutic agent to treat, ameliorate, or prevent a desired disease or condition, or to exhibit a detectable therapeutic or preventative effect. The effect can be detected by, for example, chemical markers or antigen levels. Therapeutic effects also include reduction in physical symptoms, such as decreased body temperature, when given to a patient that is febrile. The precise effective amount for a subject will depend upon the subject's size and health, the nature and extent of the condition, and the therapeutics or combination of therapeutics selected for administration. Thus, it is not useful to specify an exact effective amount in advance. However, the effective amount for a given situation can be determined by routine experimentation and is within the judgment of the clinician.

For purposes of the present invention, an effective dose will be from about 0.01 mg/ + kg to 50 mg/kg or 0.05 mg/kg to about 10 mg/kg of the DNA constructs in the individual to which it is administered.

A pharmaceutical composition can also contain a pharmaceutically acceptable carrier. The term "pharmaceutically acceptable carrier" refers to a carrier for administration of a therapeutic agent, such as antibodies or a polypeptide, genes, and other therapeutic agents. The term refers to any pharmaceutical carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition, and which may be administered without undue toxicity. Suitable carriers may be large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, and inactive virus particles. Such carriers are well known to those of ordinary skill in the art.

Pharmaceutically acceptable salts can be used therein, for example, mineral acid salts such as hydrochlorides, hydrobromides, phosphates, sulfates, and the like; and the salts of

organic acids such as acetates, propionates, malonates, benzoates, and the like. A thorough discussion of pharmaceutically acceptable excipients is available in Remington's Pharmaceutical Sciences (Mack Pub. Co., N.J. 1991).

Pharmaceutically acceptable carriers in therapeutic compositions may contain liquids such as water, saline, glycerol and ethanol. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles. Typically, the therapeutic compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared. Liposomes are included within the definition of a pharmaceutically acceptable carrier.

Delivery Methods

Once formulated, the compositions of the invention can be administered directly to the subject. The subjects to be treated can be animals; in particular, human subjects can be treated.

Direct delivery of the compositions will generally be accomplished by injection, either subcutaneously, intraperitoneally, intravenously or intramuscularly or delivered to the interstitial space of a tissue. The compositions can also be administered into a lesion. Other modes of administration include oral and pulmonary administration, suppositories, and transdermal and transcutaneous applications, needles, and gene guns or hyposprays. Dosage treatment may be a single dose schedule or a multiple dose schedule.

Vaccines

Vaccines according to the invention may either be prophylactic (i.e., to prevent infection) or therapeutic (i.e., to treat disease after infection).

Such vaccines comprise immunizing antigen(s) or immunogen(s), immunogenic polypeptide, protein(s) or protein fragments, or nucleic acids (e.g., ribonucleic acid or deoxyribonucleic acid), usually in combination with "pharmaceutically acceptable carriers," which include any carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition. Suitable carriers are typically large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, lipid aggregates (such as oil droplets or

liposomes), and inactive virus particles. Such carriers are well known to those of ordinary skill in the art. Additionally, these carriers may function as immunostimulating agents ("adjuvants"). Furthermore, the immunogen or antigen may be conjugated to a bacterial toxoid, such as a toxoid from diphtheria, tetanus, cholera, *H. pylori*, etc. pathogens.

Preferred adjuvants to enhance effectiveness of the composition include, but are not limited to: (1) aluminum salts (alum), such as aluminum hydroxide, aluminum phosphate, aluminum sulfate, etc; (2) oil-in-water emulsion formulations (with or without other specific immunostimulating agents such as muramyl peptides (see below) or bacterial cell wall components), such as for example (a) MF59 (PCT Publ. No. WO 90/14837), containing 5% Squalene, 0.5% Tween 80, and 0.5% Span 85 (optionally containing various amounts of MTP-PE (see below), although not required) formulated into submicron particles using a microfluidizer such as Model 110Y microfluidizer (Microfluidics, Newton, MA), (b) SAF, containing 10% Squalane, 0.4% Tween 80, 5% pluronic-blocked polymer L121, and thr-MDP (see below) either microfluidized into a submicron emulsion or vortexed to generate a larger particle size emulsion, and (c) RibiTM adjuvant system (RAS), (Ribi Immunochem, Hamilton, MT) containing 2% Squalene, 0.2% Tween 80, and one or more bacterial cell wall components from the group consisting of monophosphorylipid A (MPL), trehalose dimycolate (TDM), and cell wall skeleton (CWS), preferably MPL + CWS (DetoxTM); (3) saponin adjuvants, such as StimulonTM (Cambridge Bioscience, Worcester, MA) may be used or particles generated therefrom such as ISCOMs (immunostimulating complexes); (4) Complete Freund's Adjuvant (CFA) and Incomplete Freund's Adjuvant (IFA); (5) cytokines, such as interleukins (e.g., IL-1, IL-2, IL-4, IL-5, IL-6, IL-7, IL-12, etc.). interferons (e.g., gamma interferon), macrophage colony stimulating factor (M-CSF), tumor necrosis factor (TNF), etc; (6) detoxified mutants of a bacterial ADP-ribosylating toxin such as a cholera toxin (CT), a pertussis toxin (PT), or an E. coli heat-labile toxin (LT), particularly LT-K63, LT-R72, CT-S109, PT-K9/G129; see, e.g., WO 93/13302 and WO 92/19265; and (7) other substances that act as immunostimulating agents to enhance the effectiveness of the composition. Alum and MF59 are preferred.

As mentioned above, muramyl peptides include, but are not limited to, N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acetyl-normuramyl-L-alanyl-D-isoglutamine (nor-MDP), N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-sn-glycero-3-huydroxyphosphoryloxy)-ethylamine (MTP-PE), etc.

The vaccine compositions comprising immunogenic compositions (e.g., which may include the antigen, pharmaceutically acceptable carrier, and adjuvant) typically will contain diluents, such as water, saline, glycerol, ethanol, etc. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles. Alternatively, vaccine compositions comprising immunogenic compositions may comprise an antigen, polypeptide, protein, protein fragment or nucleic acid in a pharmaceutically acceptable carrier.

More specifically, vaccines comprising immunogenic compositions comprise an immunologically effective amount of the immunogenic polypeptides, as well as any other of the above-mentioned components, as needed. By "immunologically effective amount", it is meant that the administration of that amount to an individual, either in a single dose or as part of a series, is effective for treatment or prevention. This amount varies depending upon the health and physical condition of the individual to be treated, the taxonomic group of individual to be treated (e.g., nonhuman primate, primate, etc.), the capacity of the individual's immune system to synthesize antibodies, the degree of protection desired, the formulation of the vaccine, the treating doctor's assessment of the medical situation, and other relevant factors. It is expected that the amount will fall in a relatively broad range that can be determined through routine trials.

Typically, the vaccine compositions or immunogenic compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared. The preparation also may be emulsified or encapsulated in liposomes for enhanced adjuvant effect, as discussed above under pharmaceutically acceptable carriers.

The immunogenic compositions are conventionally administered parenterally, e.g., by injection, either subcutaneously or intramuscularly. Additional formulations suitable for other modes of administration include oral and pulmonary formulations, suppositories, and transdermal and transcutaneous applications. Dosage treatment may be a single dose schedule or a multiple dose schedule. The vaccine may be administered in conjunction with other immunoregulatory agents.

As an alternative to protein-based vaccines, DNA vaccination may be employed (e.g., Robinson & Torres (1997) Seminars in Immunology 9:271-283; Donnelly et al. (1997) Annu Rev Immunol 15:617-648).

Gene Delivery Vehicles

Gene therapy vehicles for delivery of constructs, including a coding sequence of a therapeutic of the invention, to be delivered to the mammal for expression in the mammal, can be administered either locally or systemically. These constructs can utilize viral or non-viral vector approaches in *in vivo* or *ex vivo* modality. Expression of such coding sequence can be induced using endogenous mammalian or heterologous promoters. Expression of the coding sequence in vivo can be either constitutive or regulated.

The invention includes gene delivery vehicles capable of expressing the contemplated nucleic acid sequences. The gene delivery vehicle is preferably a viral vector and, more preferably, a retroviral, adenoviral, adeno-associated viral (AAV), herpes viral, or alphavirus vector. The viral vector can also be an astrovirus, coronavirus, orthomyxovirus, papovavirus, paramyxovirus, parvovirus, picornavirus, poxvirus, or togavirus viral vector. See generally, Jolly (1994) Cancer Gene Therapy 1:51-64; Kimura (1994) Human Gene Therapy 5:845-852; Connelly (1995) Human Gene Therapy 6:185-193; and Kaplitt (1994) Nature Genetics 6:148-153.

Retroviral vectors are well known in the art, including B, C and D type retroviruses, xenotropic retroviruses (for example, NZB-X1, NZB-X2 and NZB9-1 (see O'Neill (1985) J. Virol. 53:160) polytropic retroviruses e.g., MCF and MCF-MLV (see Kelly (1983) J. Virol. 45:291), spumaviruses and lentiviruses. See RNA Tumor Viruses, Second Edition, Cold. Spring Harbor Laboratory, 1985.

Portions of the retroviral gene therapy vector may be derived from different retroviruses. For example, retrovector LTRs may be derived from a Murine Sarcoma Virus, a tRNA binding site from a Rous Sarcoma Virus, a packaging signal from a Murine Leukemia Virus, and an origin of second strand synthesis from an Avian Leukosis Virus.

These recombinant retroviral vectors may be used to generate transduction competent retroviral vector particles by introducing them into appropriate packaging cell lines (see US patent 5,591,624). Retrovirus vectors can be constructed for site-specific integration into host cell DNA by incorporation of a chimeric integrase enzyme into the retroviral particle (see WO96/37626). It is preferable that the recombinant viral vector is a replication defective recombinant virus.

Packaging cell lines suitable for use with the above-described retrovirus vectors are well known in the art, are readily prepared (see WO95/30763 and WO92/05266), and can be used to create producer cell lines (also termed vector cell lines or "VCLs") for the production of recombinant vector particles. Preferably, the packaging cell lines are made from human parent cells (e.g., HT1080 cells) or mink parent cell lines, which eliminates inactivation in human serum.

Preferred retroviruses for the construction of retroviral gene therapy vectors include Avian Leukosis Virus, Bovine Leukemia, Virus, Murine Leukemia Virus, Mink-Cell Focus-Inducing Virus, Murine Sarcoma Virus, Reticuloendotheliosis Virus and Rous Sarcoma Virus. Particularly preferred Murine Leukemia Viruses include 4070A and 1504A (Hartley and Rowe (1976) *J Virol* 19:19-25), Abelson (ATCC No. VR-999), Friend (ATCC No. VR-245), Graffi, Gross (ATCC Nol VR-590), Kirsten, Harvey Sarcoma Virus and Rauscher (ATCC No. VR-998) and Moloney Murine Leukemia Virus (ATCC No. VR-190). Such retroviruses may be obtained from depositories or collections such as the American Type Culture Collection ("ATCC") in Rockville, Maryland or isolated from known sources using commonly available techniques.

Exemplary known retroviral gene therapy vectors employable in this invention include those described in patent applications GB2200651, EP0415731, EP0345242, EP0334301, WO89/02468; WO89/05349, WO89/09271, WO90/02806, WO90/07936, WO94/03622, WO93/25698, WO93/25234, WO93/11230, WO93/10218, WO91/02805, WO91/02825, WO95/07994, US 5,219,740, US 4,405,712, US 4,861,719, US 4,980,289, US 4,777,127, US 5,591,624. See also Vile (1993) Cancer Res 53:3860-3864; Vile (1993) Cancer Res 53:962-967; Ram (1993) Cancer Res 53 (1993) 83-88; Takamiya (1992) J Neurosci Res 33:493-503; Baba (1993) J Neurosurg 79:729-735; Mann (1983) Cell 33:153; Cane (1984) Proc Natl Acad Sci 81:6349; and Miller (1990) Human Gene Therapy 1.

Human adenoviral gene therapy vectors are also known in the art and employable in this invention. See, for example, Berkner (1988) *Biotechniques* 6:616 and Rosenfeld (1991) *Science* 252:431, and WO93/07283, WO93/06223, and WO93/07282. Exemplary known adenoviral gene therapy vectors employable in this invention include those described in the above referenced documents and in WO94/12649, WO93/03769, WO93/19191, WO94/28938, WO95/11984, WO95/00655, WO95/27071, WO95/29993, WO95/34671, WO96/05320, WO94/08026, WO94/11506, WO93/06223, WO94/24299, WO95/14102,

I STAR WALL THE START

WO95/24297, WO95/02697, WO94/28152, WO94/24299, WO95/09241, WO95/25807, WO95/05835, WO94/18922 and WO95/09654. Alternatively, administration of DNA linked to killed adenovirus as described in Curiel (1992) Hum. Gene Ther. 3:147-154 may be employed. The gene delivery vehicles of the invention also include adenovirus associated virus (AAV) vectors. Leading and preferred examples of such vectors for use in this invention are the AAV-2 based vectors disclosed in Srivastava, WO93/09239. Most preferred AAV vectors comprise the two AAV inverted terminal repeats in which the native D-sequences are modified by substitution of nucleotides, such that at least 5 native nucleotides and up to 18 native nucleotides, preferably at least 10 native nucleotides up to 18 native nucleotides, most preferably 10 native nucleotides are retained and the remaining nucleotides of the D-sequence are deleted or replaced with non-native nucleotides. The native D-sequences of the AAV inverted terminal repeats are sequences of 20 consecutive nucleotides in each AAV inverted terminal repeat (i.e., there is one sequence at each end) which are not involved in HP formation. The non-native replacement nucleotide may be any nucleotide other than the nucleotide found in the native D-sequence in the same position. Other employable exemplary AAV vectors are pWP-19, pWN-1, both of which are disclosed in Nahreini (1993) Gene 124:257-262. Another example of such an AAV vector is psub201 (see Samulski (1987) J. Virol. 61:3096). Another exemplary AAV vector is the Double-D ITR vector. Construction of the Double-D ITR vector is disclosed in US Patent 5,478,745. Still other vectors are those disclosed in Carter US Patent 4,797,368 and Muzyczka US Patent 5,139,941, Chartejee US Patent 5,474,935, and Kotin WO94/288157. Yet a further example of an AAV vector employable in this invention is SSV9AFABTKneo, which contains the AFP enhancer and albumin promoter and directs expression predominantly in the liver. Its structure and construction are disclosed in Su (1996) Human Gene Therapy 7:463-470. Additional AAV gene therapy vectors are described in US 5,354,678, US 5,173,414, US 5,139,941, and US 5,252,479.

The gene therapy vectors comprising sequences of the invention also include herpes vectors. Leading and preferred examples are herpes simplex virus vectors containing a sequence encoding a thymidine kinase polypeptide such as those disclosed in US 5,288,641 and EP0176170 (Roizman). Additional exemplary herpes simplex virus vectors include HFEM/ICP6-LacZ disclosed in WO95/04139 (Wistar Institute), pHSVlac described in Geller (1988) Science 241:1667-1669 and in WO90/09441 and WO92/07945, HSV Us3::pgC-lacZ

described in Fink (1992) *Human Gene Therapy* 3:11-19 and HSV 7134, 2 RH 105 and GAL4 described in EP 0453242 (Breakefield), and those deposited with the ATCC as accession numbers ATCC VR-977 and ATCC VR-260.

Also contemplated are alpha virus gene therapy vectors that can be employed in this invention. Preferred alpha virus vectors are Sindbis viruses vectors. Togaviruses, Semliki Forest virus (ATCC VR-67; ATCC VR-1247), Middleberg virus (ATCC VR-370), Ross River virus (ATCC VR-373; ATCC VR-1246), Venezuelan equine encephalitis virus (ATCC VR923; ATCC VR-1250; ATCC VR-1249; ATCC VR-532), and those described in US patents 5,091,309, 5,217,879, and WO92/10578. More particularly, those alpha virus vectors described in U.S. Serial No. 08/405,627, filed March 15, 1995,WO94/21792, WO92/10578, WO95/07994, US 5,091,309 and US 5,217,879 are employable. Such alpha viruses may be obtained from depositories or collections such as the ATCC in Rockville, Maryland or isolated from known sources using commonly available techniques. Preferably, alphavirus vectors with reduced cytotoxicity are used (see USSN 08/679640).

DNA vector systems such as eukarytic layered expression systems are also useful for expressing the nucleic acids of the invention. SeeWO95/07994 for a detailed description of eukaryotic layered expression systems. Preferably, the eukaryotic layered expression systems of the invention are derived from alphavirus vectors and most preferably from Sindbis viral vectors.

Other viral vectors suitable for use in the present invention include those derived from poliovirus, for example ATCC VR-58 and those described in Evans, Nature 339 (1989) 385 and Sabin (1973) *J. Biol. Standardization* 1:115; rhinovirus, for example ATCC VR-1110 and those described in Arnold (1990) *J Cell Biochem* L401; pox viruses such as canary pox virus or vaccinia virus, for example ATCC VR-111 and ATCC VR-2010 and those described in Fisher-Hoch (1989) *Proc Natl Acad Sci* 86:317; Flexner (1989) *Ann NY Acad Sci* 569:86, Flexner (1990) *Vaccine* 8:17; in US 4,603,112 and US 4,769,330 and WO89/01973; SV40 virus, for example ATCC VR-305 and those described in Mulligan (1979) *Nature* 277:108 and Madzak (1992) *J Gen Virol* 73:1533; influenza virus, for example ATCC VR-797 and recombinant influenza viruses made employing reverse genetics techniques as described in US 5,166,057 and in Enami (1990) *Proc Natl Acad Sci* 87:3802-3805; Enami & Palese (1991) *J Virol* 65:2711-2713 and Luytjes (1989) *Cell* 59:110, (see also McMichael (1983) *NEJ Med* 309:13, and Yap (1978) *Nature* 273:238 and Nature (1979) 277:108); human

immunodeficiency virus as described in EP-0386882 and in Buchschacher (1992) *J. Virol.* 66:2731; measles virus, for example ATCC VR-67 and VR-1247 and those described in EP-0440219; Aura virus, for example ATCC VR-368; Bebaru virus, for example ATCC VR-600 and ATCC VR-1240; Cabassou virus, for example ATCC VR-922; Chikungunya virus, for example ATCC VR-64 and ATCC VR-1241; Fort Morgan Virus, for example ATCC VR-924; Getah virus, for example ATCC VR-369 and ATCC VR-1243; Kyzylagach virus, for example ATCC VR-927; Mayaro virus, for example ATCC VR-66; Mucambo virus, for example ATCC VR-580 and ATCC VR-1244; Ndumu virus, for example ATCC VR-371; Pixuna virus, for example ATCC VR-372 and ATCC VR-1245; Tonate virus, for example ATCC VR-925; Triniti virus, for example ATCC VR-469; Una virus, for example ATCC VR-374; Whataroa virus, for example ATCC VR-926; Y-62-33 virus, for example ATCC VR-375; O'Nyong virus, Eastern encephalitis virus, for example ATCC VR-1251, ATCC VR-1242; Western encephalitis virus, for example ATCC VR-70, ATCC VR-1251, ATCC VR-622 and ATCC VR-1252; and coronavirus, for example ATCC VR-740 and those described in Hamre (1966) *Proc Soc Exp Biol Med* 121:190.

Delivery of the compositions of this invention into cells is not limited to the above mentioned viral vectors. Other delivery methods and media may be employed such as, for example, nucleic acid expression vectors, polycationic condensed DNA linked or unlinked to killed adenovirus alone, for example see US Serial No. 08/366,787, filed December 30, 1994 and Curiel (1992) *Hum Gene Ther* 3:147-154 ligand linked DNA, for example see Wu (1989) *J Biol Chem* 264:16985-16987, eucaryotic cell delivery vehicles cells, for example see US Serial No.08/240,030, filed May 9, 1994, and US Serial No. 08/404,796, deposition of photopolymerized hydrogel materials, hand-held gene transfer particle gun, as described in US Patent 5,149,655, ionizing radiation as described in US5,206,152 and in WO92/11033, nucleic charge neutralization or fusion with cell membranes. Additional approaches are described in Philip (1994) *Mol Cell Biol* 14:2411-2418 and in Woffendin (1994) *Proc Natl Acad Sci* 91:1581-1585.

Particle mediated gene transfer may be employed, for example see US Serial No. 60/023,867. Briefly, the sequence can be inserted into conventional vectors that contain conventional control sequences for high level expression, and then incubated with synthetic gene transfer molecules such as polymeric DNA-binding cations like polylysine, protamine, and albumin, linked to cell targeting ligands such as asialoorosomucoid, as described in Wu &

Wu (1987) J. Biol. Chem. 262:4429-4432, insulin as described in Hucked (1990) Biochem Pharmacol 40:253-263, galactose as described in Plank (1992) Bioconjugate Chem 3:533-539, lactose or transferrin.

Naked DNA may also be employed to transform a host cell. Exemplary naked DNA introduction methods are described in WO 90/11092 and US 5,580,859. Uptake efficiency may be improved using biodegradable latex beads. DNA coated latex beads are efficiently transported into cells after endocytosis initiation by the beads. The method may be improved further by treatment of the beads to increase hydrophobicity and thereby facilitate disruption of the endosome and release of the DNA into the cytoplasm.

Liposomes that can act as gene delivery vehicles are described in U.S. 5,422,120, WO95/13796, WO94/23697, WO91/14445 and EP-524,968. As described in USSN. 60/023,867, on non-viral delivery, the nucleic acid sequences encoding a polypeptide can be inserted into conventional vectors that contain conventional control sequences for high level expression, and then be incubated with synthetic gene transfer molecules such as polymeric DNA-binding cations like polylysine, protamine, and albumin, linked to cell targeting ligands such as asialoorosomucoid, insulin, galactose, lactose, or transferrin. Other delivery systems include the use of liposomes to encapsulate DNA comprising the gene under the control of a variety of tissue-specific or ubiquitously-active promoters. Further non-viral delivery suitable for use includes mechanical delivery systems such as the approach described in Woffendin et al (1994) Proc. Natl. Acad. Sci. USA 91(24):11581-11585. Moreover, the coding sequence and the product of expression of such can be delivered through deposition of photopolymerized hydrogel materials. Other conventional methods for gene delivery that can be used for delivery of the coding sequence include, for example, use of hand-held gene transfer particle gun, as described in U.S. 5,149,655; use of ionizing radiation for activating transferred gene, as described in U.S. 5,206,152 and WO92/11033.

Exemplary liposome and polycationic gene delivery vehicles are those described in US 5,422,120 and 4,762,915; inWO 95/13796; WO94/23697; and WO91/14445; in EP-0524968; and in Stryer, Biochemistry, pages 236-240 (1975) W.H. Freeman, San Francisco; Szoka (1980) Biochem Biophys Acta 600:1; Bayer (1979) Biochem Biophys Acta 550:464; Rivnay (1987) Meth Enzymol 149:119; Wang (1987) Proc Natl Acad Sci 84:7851; Plant (1989) Anal Biochem 176:420.

A polynucleotide composition can comprises therapeutically effective amount of a gene therapy vehicle, as the term is defined above. For purposes of the present invention, an effective dose will be from about 0.01 mg/kg to 50 mg/kg or 0.05 mg/kg to about 10 mg/kg of the DNA constructs in the individual to which it is administered.

Delivery Methods

Once formulated, the polynucleotide compositions of the invention can be administered (1) directly to the subject; (2) delivered ex vivo, to cells derived from the subject; or (3) in vitro for expression of recombinant proteins. The subjects to be treated can be mammals or birds. Also, human subjects can be treated.

Direct delivery of the compositions will generally be accomplished by injection, either subcutaneously, intraperitoneally, intravenously or intramuscularly or delivered to the interstitial space of a tissue. The compositions can also be administered into a tumor or lesion. Other modes of administration include oral and pulmonary administration, suppositories, and transdermal applications, needles, and gene guns or hyposprays. Dosage treatment may be a single dose schedule or a multiple dose schedule.

Methods for the *ex vivo* delivery and reimplantation of transformed cells into a subject are known in the art and described in *eg.* WO93/14778. Examples of cells useful in ex vivo applications include, for example, stem cells, particularly hematopoetic, lymph cells, macrophages, dendritic cells, or tumor cells.

Generally, delivery of nucleic acids for both ex vivo and in vitro applications can be accomplished by the following procedures, for example, dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei, all well known in the art.

Polynucleotide and polypeptide pharmaceutical compositions

In addition to the pharmaceutically acceptable carriers and salts described above, the following additional agents can be used with polynucleotide and/or polypeptide compositions.

A.Polypeptides

One example are polypeptides which include, without limitation: asioloorosomucoid (ASOR); transferrin; asialoglycoproteins; antibodies; antibody fragments; ferritin; interleukins; interferons, granulocyte, macrophage colony stimulating factor (GM-CSF),

granulocyte colony stimulating factor (G-CSF), macrophage colony stimulating factor (M-CSF), stem cell factor and erythropoietin. Viral antigens, such as envelope proteins, can also be used. Also, proteins from other invasive organisms, such as the 17 amino acid peptide from the circumsporozoite protein of plasmodium falciparum known as RII. B.Hormones, Vitamins, Etc.

Other groups that can be included are, for example: hormones, steroids, androgens, estrogens, thyroid hormone, or vitamins, folic acid.

C.Polyalkylenes, Polysaccharides, etc.

Also, polyalkylene glycol can be included with the desired polynucleotides or polypeptides. In a preferred embodiment, the polyalkylene glycol is polyethlylene glycol. In addition, mono-, di-, or polysaccarides can be included. In a preferred embodiment of this aspect, the polysaccharide is dextran or DEAE-dextran. Also, chitosan and poly(lactide-co-glycolide)

D.Lipids, and Liposomes

The desired polynucleotide or polypeptide can also be encapsulated in lipids or packaged in liposomes prior to delivery to the subject or to cells derived therefrom.

Lipid encapsulation is generally accomplished using liposomes which are able to stably bind or entrap and retain nucleic acid. The ratio of condensed polynucleotide or polypeptide to lipid preparation can vary but will generally be around 1:1 (mg DNA:micromoles lipid), or more of lipid. For a review of the use of liposomes as carriers for delivery of nucleic acids, see, Hug and Sleight (1991) *Biochim. Biophys. Acta.* 1097:1-17; Straubinger (1983) *Meth. Enzymol.* 101:512-527.

Liposomal preparations for use in the present invention include cationic (positively charged), anionic (negatively charged) and neutral preparations. Cationic liposomes have been shown to mediate intracellular delivery of plasmid DNA (Felgner (1987) *Proc. Natl. Acad. Sci. USA* 84:7413-7416); mRNA (Malone (1989) *Proc. Natl. Acad. Sci. USA* 86:6077-6081); and purified transcription factors (Debs (1990) *J. Biol. Chem.* 265:10189-10192), in functional form.

Cationic liposomes are readily available. For example, N[1-2,3-dioleyloxy)propyl]-N,N,N-triethylammonium (DOTMA) liposomes are available under the trademark Lipofectin, from GIBCO BRL, Grand Island, NY. (See, also, Felgner supra). Other commercially available liposomes include transfectace (DDAB/DOPE) and

DOTAP/DOPE (Boerhinger). Other cationic liposomes can be prepared from readily available materials using techniques well known in the art. See, eg. Szoka (1978) Proc. Natl. Acad. Sci. USA 75:4194-4198; WO90/11092 for a description of the synthesis of DOTAP (1,2-bis(oleoyloxy)-3-(trimethylammonio)propane) liposomes.

Similarly, anionic and neutral liposomes are readily available, such as from Avanti Polar Lipids (Birmingham, AL), or can be easily prepared using readily available materials. Such materials include phosphatidyl choline, cholesterol, phosphatidyl ethanolamine, dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), dioleoylphoshatidyl ethanolamine (DOPE), among others. These materials can also be mixed with the DOTMA and DOTAP starting materials in appropriate ratios. Methods for making liposomes using these materials are well known in the art.

The liposomes can comprise multilammelar vesicles (MLVs), small unilamellar vesicles (SUVs), or large unilamellar vesicles (LUVs). The various liposome-nucleic acid complexes are prepared using methods known in the art. See eg. Straubinger (1983) Meth. Immunol. 101:512-527; Szoka (1978) Proc. Natl. Acad. Sci. USA 75:4194-4198; Papahadjopoulos (1975) Biochim. Biophys. Acta 394:483; Wilson (1979) Cell 17:77); Deamer & Bangham (1976) Biochim. Biophys. Acta 443:629; Ostro (1977) Biochem. Biophys. Res. Commun. 76:836; Fraley (1979) Proc. Natl. Acad. Sci. USA 76:3348); Enoch & Strittmatter (1979) Proc. Natl. Acad. Sci. USA 76:145; Fraley (1980) J. Biol. Chem. (1980) 255:10431; Szoka & Papahadjopoulos (1978) Proc. Natl. Acad. Sci. USA 75:145; and Schaefer-Ridder (1982) Science 215:166.

E.Lipoproteins

In addition, lipoproteins can be included with the polynucleotide or polypeptide to be delivered. Examples of lipoproteins to be utilized include: chylomicrons, HDL, IDL, LDL, and VLDL. Mutants, fragments, or fusions of these proteins can also be used. Also, modifications of naturally occurring lipoproteins can be used, such as acetylated LDL. These lipoproteins can target the delivery of polynucleotides to cells expressing lipoprotein receptors. Preferably, if lipoproteins are including with the polynucleotide to be delivered, no other targeting ligand is included in the composition.

Naturally occurring lipoproteins comprise a lipid and a protein portion. The protein portion are known as apoproteins. At the present, apoproteins A, B, C, D, and E have been

isolated and identified. At least two of these contain several proteins, designated by Roman numerals, AI, AII, AIV; CI, CII, CIII.

A lipoprotein can comprise more than one apoprotein. For example, naturally occurring chylomicrons comprises of A, B, C, and E, over time these lipoproteins lose A and acquire C and E apoproteins. VLDL comprises A, B, C, and E apoproteins, LDL comprises apoprotein B; and HDL comprises apoproteins A, C, and E.

The amino acid of these apoproteins are known and are described in, for example, Breslow (1985) Annu Rev. Biochem 54:699; Law (1986) Adv. Exp Med. Biol. 151:162; Chen (1986) J Biol Chem 261:12918; Kane (1980) Proc Natl Acad Sci USA 77:2465; and Utermann (1984) Hum Genet 65:232.

Lipoproteins contain a variety of lipids including, triglycerides, cholesterol (free and esters), and phopholipids. The composition of the lipids varies in naturally occurring lipoproteins. For example, chylomicrons comprise mainly triglycerides. A more detailed description of the lipid content of naturally occurring lipoproteins can be found, for example, in *Meth. Enzymol.* 128 (1986). The composition of the lipids are chosen to aid in conformation of the apoprotein for receptor binding activity. The composition of lipids can also be chosen to facilitate hydrophobic interaction and association with the polynucleotide binding molecule.

Naturally occurring lipoproteins can be isolated from serum by ultracentrifugation, for instance. Such methods are described in *Meth. Enzymol.* (supra); Pitas (1980) J. Biochem. 255:5454-5460 and Mahey (1979) J Clin. Invest 64:743-750.

Lipoproteins can also be produced by *in vitro* or recombinant methods by expression of the apoprotein genes in a desired host cell. See, for example, Atkinson (1986) *Annu Rev Biophys Chem* 15:403 and Radding (1958) *Biochim Biophys Acta* 30: 443.

Lipoproteins can also be purchased from commercial suppliers, such as Biomedical Techniologies, Inc., Stoughton, Massachusetts, USA.

Further description of lipoproteins can be found in Zuckermann et al., PCT. Appln. No. US97/14465.

F.Polycationic Agents

Polycationic agents can be included, with or without lipoprotein, in a composition with the desired polynucleotide or polypeptide to be delivered.

25

Polycationic agents, typically, exhibit a net positive charge at physiological relevant pH and are capable of neutralizing the electrical charge of nucleic acids to facilitate delivery to a desired location. These agents have both in vitro, ex vivo, and in vivo applications. Polycationic agents can be used to deliver nucleic acids to a living subject either intramuscularly, subcutaneously, etc.

The following are examples of useful polypeptides as polycationic agents: polylysine, polyarginine, polyornithine, and protamine. Other examples include histones, protamines, human serum albumin, DNA binding proteins, non-histone chromosomal proteins, coat proteins from DNA viruses, such as (X174, transcriptional factors also contain domains that bind DNA and therefore may be useful as nucleic aid condensing agents. Briefly, transcriptional factors such as C/CEBP, c-jun, c-fos, AP-1, AP-2, AP-3, CPF, Prot-1, Sp-1, Oct-1, Oct-2, CREP, and TFIID contain basic domains that bind DNA sequences.

Organic polycationic agents include: spermine, spermidine, and purtrescine.

The dimensions and of the physical properties of a polycationic agent can be extrapolated from the list above, to construct other polypeptide polycationic agents or to produce synthetic polycationic agents.

Synthetic Polycationic Agents

Synthetic polycationic agents which are useful include, for example, DEAE-dextran, polybrene. Lipofectin , and lipofectAMINE are monomers that form polycationic complexes when combined with polynucleotides or polypeptides.

Immunodiagnostic Assays

Neisserial antigens of the invention can be used in immunoassays to detect antibody levels (or, conversely, anti-Neisserial antibodies can be used to detect antigen levels). Immunoassays based on well defined, recombinant antigens can be developed to replace invasive diagnostics methods. Antibodies to Neisserial proteins within biological samples, including for example, blood or serum samples, can be detected. Design of the immunoassays is subject to a great deal of variation, and a variety of these are known in the art. Protocols for the immunoassay may be based, for example, upon competition, or direct reaction, or sandwich type assays. Protocols may also, for example, use solid supports, or may be by immunoprecipitation. Most assays involve the use of labeled antibody or polypeptide; the labels may be, for example, fluorescent, chemiluminescent, radioactive, or dye molecules. Assays which amplify the signals from the probe are also known; examples of which are

assays which utilize biotin and avidin, and enzyme-labeled and mediated immunoassays, such as ELISA assays.

Kits suitable for immunodiagnosis and containing the appropriate labeled reagents are constructed by packaging the appropriate materials, including the compositions of the invention, in suitable containers, along with the remaining reagents and materials (for example, suitable buffers, salt solutions, etc.) required for the conduct of the assay, as well as suitable set of assay instructions.

Nucleic Acid Hybridisation

"Hybridization" refers to the association of two nucleic acid sequences to one another by hydrogen bonding. Typically, one sequence will be fixed to a solid support and the other will be free in solution. Then, the two sequences will be placed in contact with one another under conditions that favor hydrogen bonding. Factors that affect this bonding include: the type and volume of solvent; reaction temperature; time of hybridization; agitation; agents to block the non-specific attachment of the liquid phase sequence to the solid support (Denhardt's reagent or BLOTTO); concentration of the sequences; use of compounds to increase the rate of association of sequences (dextran sulfate or polyethylene glycol); and the stringency of the washing conditions following hybridization. See Sambrook *et al.* [supra] Volume 2, chapter 9, pages 9.47 to 9.57.

"Stringency" refers to conditions in a hybridization reaction that favor association of very similar sequences over sequences that differ. For example, the combination of temperature and salt concentration should be chosen that is approximately 120 to 200 \Box C below the calculated Tm of the hybrid under study. The temperature and salt conditions can often be determined empirically in preliminary experiments in which samples of genomic DNA immobilized on filters are hybridized to the sequence of interest and then washed under conditions of different stringencies. See Sambrook *et al.* at page 9.50.

Variables to consider when performing, for example, a Southern blot are (1) the complexity of the DNA being blotted and (2) the homology between the probe and the sequences being detected. The total amount of the fragment(s) to be studied can vary a magnitude of 10, from 0.1 to 1µg for a plasmid or phage digest to 10^{-9} to 10^{-8} g for a single copy gene in a highly complex eukaryotic genome. For lower complexity polynucleotides, substantially shorter blotting, hybridization, and exposure times, a smaller amount of starting polynucleotides, and lower specific activity of probes can be used. For example, a single-copy

yeast gene can be detected with an exposure time of only 1 hour starting with 1 μ g of yeast DNA, blotting for two hours, and hybridizing for 4-8 hours with a probe of 10^8 cpm/ μ g. For a single-copy mammalian gene a conservative approach would start with 10 μ g of DNA, blot overnight, and hybridize overnight in the presence of 10% dextran sulfate using a probe of greater than 10^8 cpm/ μ g, resulting in an exposure time of ~24 hours.

Several factors can affect the melting temperature (Tm) of a DNA-DNA hybrid between the probe and the fragment of interest, and consequently, the appropriate conditions for hybridization and washing. In many cases the probe is not 100% homologous to the fragment. Other commonly encountered variables include the length and total G+C content of the hybridizing sequences and the ionic strength and formamide content of the hybridization buffer. The effects of all of these factors can be approximated by a single equation:

Tm= $81 + 16.6(\log_{10}\text{Ci}) + 0.4[\%(G + C)]-0.6(\%\text{formamide}) - 600/n-1.5(\%\text{mismatch}).$ where Ci is the salt concentration (monovalent ions) and n is the length of the hybrid in base pairs (slightly modified from Meinkoth & Wahl (1984) Anal. Biochem. 138: 267-284).

In designing a hybridization experiment, some factors affecting nucleic acid hybridization can be conveniently altered. The temperature of the hybridization and washes and the salt concentration during the washes are the simplest to adjust. As the temperature of the hybridization increases (*ie.* stringency), it becomes less likely for hybridization to occur between strands that are nonhomologous, and as a result, background decreases. If the radiolabeled probe is not completely homologous with the immobilized fragment (as is frequently the case in gene family and interspecies hybridization experiments), the hybridization temperature must be reduced, and background will increase. The temperature of the washes affects the intensity of the hybridizing band and the degree of background in a similar manner. The stringency of the washes is also increased with decreasing salt concentrations.

In general, convenient hybridization temperatures in the presence of 50% formamide are $42\Box C$ for a probe with is 95% to 100% homologous to the target fragment, $37\Box C$ for 90% to 95% homology, and $32\Box C$ for 85% to 90% homology. For lower homologies, formamide content should be lowered and temperature adjusted accordingly, using the equation above. If the homology between the probe and the target fragment are not known, the simplest approach is to start with both hybridization and wash conditions which are nonstringent. If non-specific bands or high background are observed after autoradiography, the filter can be

washed at high stringency and reexposed. If the time required for exposure makes this approach impractical, several hybridization and/or washing stringencies should be tested in parallel.

Nucleic Acid Probe Assays

Methods such as PCR, branched DNA probe assays, or blotting techniques utilizing nucleic acid probes according to the invention can determine the presence of cDNA or mRNA. A probe is said to "hybridize" with a sequence of the invention if it can form a duplex or double stranded complex, which is stable enough to be detected.

The nucleic acid probes will hybridize to the Neisserial nucleotide sequences of the invention (including both sense and antisense strands). Though many different nucleotide sequences will encode the amino acid sequence, the native Neisserial sequence is preferred because it is the actual sequence present in cells. mRNA represents a coding sequence and so a probe should be complementary to the coding sequence; single-stranded cDNA is complementary to mRNA, and so a cDNA probe should be complementary to the non-coding sequence.

The probe sequence need not be identical to the Neisserial sequence (or its complement) — some variation in the sequence and length can lead to increased assay sensitivity if the nucleic acid probe can form a duplex with target nucleotides, which can be detected. Also, the nucleic acid probe can include additional nucleotides to stabilize the formed duplex. Additional Neisserial sequence may also be helpful as a label to detect the formed duplex. For example, a non-complementary nucleotide sequence may be attached to the 5' end of the probe, with the remainder of the probe sequence being complementary to a Neisserial sequence. Alternatively, non-complementary bases or longer sequences can be interspersed into the probe, provided that the probe sequence has sufficient complementarity with the a Neisserial sequence in order to hybridize therewith and thereby form a duplex which can be detected.

The exact length and sequence of the probe will depend on the hybridization conditions, such as temperature, salt condition and the like. For example, for diagnostic applications, depending on the complexity of the analyte sequence, the nucleic acid probe typically contains at least 10-20 nucleotides, preferably 15-25, and more preferably at least 30 nucleotides, although it may be shorter than this. Short primers generally require cooler temperatures to form sufficiently stable hybrid complexes with the template.

Probes may be produced by synthetic procedures, such as the triester method of Matteucci et al. [J. Am. Chem. Soc. (1981) 103:3185], or according to Urdea et al. [Proc. Natl. Acad. Sci. USA (1983) 80: 7461], or using commercially available automated oligonucleotide synthesizers.

The chemical nature of the probe can be selected according to preference. For certain applications, DNA or RNA are appropriate. For other applications, modifications may be incorporated eg. backbone modifications, such as phosphorothioates or methylphosphonates, can be used to increase in vivo half-life, alter RNA affinity, increase nuclease resistance etc. [eg. see Agrawal & Iyer (1995) Curr Opin Biotechnol 6:12-19; Agrawal (1996) TIBTECH 14:376-387]; analogues such as peptide nucleic acids may also be used [eg. see Corey (1997) TIBTECH 15:224-229; Buchardt et al. (1993) TIBTECH 11:384-386].

One example of a nucleotide hybridization assay is described by Urdea *et al.* in international patent application WO92/02526 [see also US patent 5,124,246].

Alternatively, the polymerase chain reaction (PCR) is another well-known means for detecting small amounts of target nucleic acids. The assay is described in: Mullis et al. [Meth. Enzymol. (1987) 155: 335-350]; US patent 4,683,195; and US patent 4,683,202. Two "primer" nucleotides hybridize with the target nucleic acids and are used to prime the reaction. The primers can comprise sequence that does not hybridize to the sequence of the amplification target (or its complement) to aid with duplex stability or, for example, to incorporate a convenient restriction site. Typically, such sequence will flank the desired Neisserial sequence.

A thermostable polymerase creates copies of target nucleic acids from the primers using the original target nucleic acids as a template. After a threshold amount of target nucleic acids are generated by the polymerase, they can be detected by more traditional methods, such as Southern blots. When using the Southern blot method, the labelled probe will hybridize to the Neisserial sequence (or its complement).

Also, mRNA or cDNA can be detected by traditional blotting techniques described in Sambrook et al [supra]. mRNA, or cDNA generated from mRNA using a polymerase enzyme, can be purified and separated using gel electrophoresis. The nucleic acids on the gel are then blotted onto a solid support, such as nitrocellulose. The solid support is exposed to a labelled probe and then washed to remove any unhybridized probe. Next, the duplexes

containing the labeled probe are detected. Typically, the probe is labelled with a radioactive moiety.

EXAMPLES

The examples describe nucleic acid sequences which have been identified in N. meningitidis, and N. gonorrhoeae along with their respective and putative translation products. Not all of the nucleic acid sequences are complete ie. they encode less than the full-length wild-type protein.

The examples are generally in the following format:

- a nucleotide sequence which has been identified in N. meningitidis
- the putative translation product of said N. meningitidis sequence
- a computer analysis of said translation product based on database comparisons
- a corresponding nucleotide sequence identified from N. gonorrhoeae
- the putative translation product of said N. gonorrhoeae sequence
- a comparision of the percentage of identity between the translation product of the N. meningitidis sequence and the N. gonorrhoeae sequence.
- a corresponding nucleotide sequence identified from strain A of N. meningitidis
- the putative translation product of said N. meningitidis strain A sequence
- a comparision of the percentage of identity between the translation product of the N. meningitidis sequence and the N. gonorrhoeae sequence.
- a description of the characteristics of the protein which indicates that it might be suitably antigenic or immunogenic.

Sequence comparisons were performed at NCBI (http://www.ncbi.nlm.nih.gov) using the algorithms BLAST, BLAST2, BLAST1, BLAST1, tBLAST1, tBLAST1, BLAST2, & tBLAST1 [eg. see also Altschul et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research 25:2289-3402]. Searches were performed against the following databases: non-redundant GenBank+EMBL+DDBJ+PDB sequences and non-redundant GenBank CDS translations+PDB+SwissProt+SPupdate+PIR sequences.

Dots within nucleotide sequences represent nucleotides which have been arbitrarily introduced in order to maintain a reading frame. In the same way, double-underlined nucleotides were removed. Lower case letters represent ambiguities which arose during

100 C 100 C 100 C 100 C

alignment of independent sequencing reactions (some of the nucleotide sequences in the examples are derived from combining the results of two or more experiments).

Nucleotide sequences were scanned in all six reading frames to predict the presence of hydrophobic domains using an algorithm based on the statistical studies of Esposti *et al.* [Critical evaluation of the hydropathy of membrane proteins (1990) *Eur J Biochem* 190:207-219]. These domains represent potential transmembrane regions or hydrophobic leader sequences.

Open reading frames were predicted from fragmented nucleotide sequences using the program ORFFINDER (NCBI).

Underlined amino acid sequences indicate possible transmembrane domains or leader sequences in the ORFs, as predicted by the PSORT algorithm (http://www.psort.nibb.ac.jp). Functional domains were also predicted using the MOTIFS program (GCG Wisconsin & PROSITE).

For each of the following examples: based on the presence of a putative leader sequence and/or several putative transmembrane domains (single-underlined) in the gonococcal protein, it is predicted that the proteins from *N. meningitidis* and *N. gonorrhoeae*, and their respective epitopes, could be useful antigens or immunogenic compositions for vaccines or diagnostics.

The standard techniques and procedures which may be employed in order to perform the invention (e.g. to utilize the disclosed sequences for vaccination or diagnostic purposes) were summarized above. This summary is not a limitation on the invention but, rather, gives examples that may be used, but are not required.

In particular, the following methods were used to express, purify and biochemically characterize the proteins of the invention.

Chromosomal DNA Preparation

N.meningitidis strain 2996 was grown to exponential phase in 100ml of GC medium, harvested by centrifugation, and resuspended in 5ml buffer (20%(w/v) Sucrose, 50mM Tris-HCl, 50mM EDTA, pH8). After 10 minutes incubation on ice, the bacteria were lysed by adding 10ml of lysis solution (50mM NaCl, 1% Na-Sarkosyl, 50µg/ml Proteinase K), and the suspension incubated at 37°C for 2 hours. Two phenol extractions (equilibrated to pH 8) and one CHCl₃/isoamylalcohol (24:1) extraction were performed. DNA was precipitated by addition of 0.3M sodium acetate and 2 volumes of ethanol, and collected by centrifugation.

BNSDOCID: <WO___9957280A2_I_>

The pellet was washed once with 70%(v/v) ethanol and redissolved in 4.0ml TE buffer (10mM Tris-HCl, 1mM EDTA, pH 8.0). The DNA concentration was measured by reading the OD at 260 nm.

Oligonucleotide design

Synthetic oligonucleotide primers were designed on the basis of the coding sequence of each ORF, using (a) the meningococcus B sequence when available, or (b) the gonococcus/meningococcus A sequence, adapted to the codon preference usage of meningococcus as necessary. Any predicted signal peptides were omitted, by designing the 5' primers to sequence immediately downstream from the predicted leader sequence.

For most ORFs, the 5' primers included two restriction enzyme recognition sites (BamHI-NdeI, BamHI-NheI, EcoRI-NdeI or EcoRI-NheI), depending on the restriction pattern of the gene of interest. The 3' primers included a XhoI or a HindIII restriction site (table 1). This procedure was established in order to direct the cloning of each amplification product (corresponding to each ORF) into two different expression systems: pGEX-KG (using BamHI-XhoI, BamHI-HindIII, EcoRI-XhoI or EcoRI-HindIII), and pET21b+ (using NdeI-XhoI, NheI-XhoI, NdeI-HindIII or NheI-HindIII).

5'-end primer tail:	CGCGGATCCCATATG	(BamHI-NdeI
	CGCGGATCCGCTAGC	(BamHI-NheI)
	CCGGAATTCTACATATG	(EcoRI-NdeI)
	CCGGAATTCTAGCTAGC	(EcoRI-NheI)
3'-end primer tail:	CCCG <u>CTCGAG</u> CCCGCTCGAG	(XhoI) (HindIII)

For cloning ORFs into the pGEX-His vector, the 5' and 3' primers contained only one restriction enzyme site (*EcoRI*, *KpnI* or *SalI* for the 5' primers and *PstI*, *XbaI*, *SphI* or *SalI* for the 3' primers). Again restriction sites were chosen according to the particular restriction pattern of the gene (table 1).

5'-end primer tail:	(AAA) AAAGAATTC	(EcoRI)
	(AAA) AAAGGTACC	(KpnI)
3'-end primer tail:	(AAA) AAA <u>CTGCAG</u>	(PstI)
	(AAA) AAA <u>TCTAGA</u>	(XbaI)

AAAGCATGC

(SphI)

5' or 3'-end primer tail:

AAAAAAGTCGAC

(SalI)

والشرائر والمرازي تباري

As well as containing the restriction enzyme recognition sequences, the primers included nucleotides which hybridized to the sequence to be amplified. The melting temperature depended on the number and type of hybridising nucleotides in the whole primer, and was determined for each primer using the formulae:

$$T_m = 4 (G+C)+ 2 (A+T)$$
 (tail excluded)
 $T_m = 64.9 + 0.41 (\% GC) - 600/N$ (whole primer)

The melting temperatures of the selected oligonucleotides were usually 65-70°C for the whole oligo and 50-55°C for the hybridising region alone.

Table 1 shows the forward and reverse primers used for each amplification. In certain cases, the sequence of the primer does not exactly match the sequence of the predicted ORF. This is because when initial amplifications were performed, the complete 5' and/or 3' sequences for some meningococcal B ORFs were not known. However the corresponding sequences had been identified in Gonococcus or in Meningoccus A. Hence, when the Meningoccus B sequence was incomplete or uncertain, Gonococcal or Meningococcal A sequences were used as the basis for primer design. These sequences were altered to take account of codon preference. It can be appreciated that, once the complete sequence is identified, this approach will no longer be necessary.

Oligonucleotides were synthesized using a Perkin Elmer 394 DNA/RNA Synthesizer, eluted from the columns in 2.0ml NH₄OH, and deprotected by 5 hours incubation at 56°C. The oligos were precipitated by addition of 0.3M Na-Acetate and 2 volumes ethanol. The samples were centrifuged and the pellets resuspended in either $100\mu l$ or 1.0ml of water. The OD₂₆₀ was determined using a Perkin Elmer Lambda Bio spectophotometer and the concentration adjusted to 2-10pmol/ μl .

Amplification

The standard PCR protocol was as follows: 50-200ng of genomic DNA was used as a template in the presence of 20-40µM of each oligonucletide primer, 400-800µM dNTPs solution, 1x PCR buffer (including 1.5mM MgCl₂), 2.5 units *TaqI* DNA polymerase (using

Perkin-Elmer AmpliTaQ, GIBCO Platinum, Pwo DNA polymerase, or Tahara Shuzo Taq polymerase). In some cases, PCR was optimsed by the addition of $10\mu l$ DMSO or $50\mu l$ 2M Betaine.

After a hot start (adding the polymerase during a preliminary 3 minute incubation of the whole mix at 95°C), each sample underwent a two-step amplification. The first 5 cycles were performed using the hybridization temperature that excluded the restriction enzyme tail of the primer (see above). This was followed by 30 cycles using the hybridization temperature calculated for the whole length oligos. The cycles were completed with a 10 minute extension step at 72°C. The standard cycles were as follows:

	Denaturation	Hybridisation	Elongation
First 5 cycles	30 seconds	30 seconds	30-60 seconds
	95°C	50-55°C	72°C
Last 30 cycles	30 seconds	30 seconds	30-60 seconds
	95°C	65-70°C	72°C

Elongation times varied according to the length of the ORF to be amplified. Amplifications were performed using either a 9600 or a 2400 Perkin Elmer GeneAmp PCR System. To check the results, 1/10 of the amplification volume was loaded onto a 1-1.5% (w/v) agarose gel and the size of each amplified fragment compared with a DNA molecular weight marker.

The amplified DNA was either loaded directly on a 1% agarose gel or first precipitated with ethanol and resuspended in a volume suitable to be loaded on a 1.0% agarose gel. The DNA fragment corresponding to the band of correct size was purified using the Qiagen Gel Extraction Kit, following the manufacturer's protocol. DNA fragments were eluted in a volume of $30\mu l$ or $50\mu l$ with either H2O or 10mM Tris, pH 8.5.

Digestion of PCR fragments

The purified DNA corresponding to the amplified fragment was doubly-digested with the appropriate restriction enzymes for; cloning into pET-21b+ and expressing the protein as a C-terminus His-tagged fusion, for cloning into pGEX-KG and expressing the protein as a N-

terminus GST-fusion, and for cloning into pGEX-His and expressing the protein as a N-terminus GST-His tagged fusion.

Each purified DNA fragment was incubated at 37°C for 3 hours to overnight with 20 units of appropriate restriction enzyme (New England Biolabs) in a volume of either 30 or 40µl in the presence of suitable digestion buffer. Digested fragments were purified using the QIAquick PCR purification kit (following the manufacturer's instructions) and eluted in a volume of 30µl or 50µl with either H2O or 10mM Tris, pH 8.5. The DNA concentration was determined by quantitative agarose gel electrophoresis (1.0% gel) in the presence of a titrated molecular weight marker.

Digestion of the cloning vectors (pET22B, pGEX-KG, pTRC-His A, pET21b+, pGEX-KG, and pGEX-His)

The vector pGEX-His is a modified pGEX-2T vector carrying a region encoding six histidine residues upstream of the thrombin cleavage site and containing the multiple cloning site of the vector pTRC99 (Pharmacia).10 μ g plasmid was double-digested with 50 units of each restriction enzyme in 200 μ l reaction volume in the presence of appropriate buffer by overnight incubation at 37°C. After loading the whole digestion on a 1% agarose gel, the band corresponding to the digested vector was purified from the gel using the Qiagen QIAquick Gel Extraction Kit and the DNA was eluted in 50 μ l of 10 mM Tris-HCl, pH 8.5. The DNA concentration was evaluated by measuring OD₂₆₀ of the sample, and adjusted to 50 μ g/ μ l. 1 μ l of plasmid was used for each cloning procedure.

 $10\mu g$ of plasmid vector was doubly-digested with 50 units of each restriction enzyme in a volume of $200\mu l$ with the appropriate buffer overnight at $37^{\circ}C$. The digest was loaded onto a 1.0% agarose gel and the band corresponding to the digested vector purified using the Qiagen QIAquick Gel Extraction Kit. DNA was eluted in $50\mu l$ of 10mM Tris-HCl, pH 8.5. The DNA concentration was evaluated by measuring OD_{260nm} and the concentration adjusted to $50\mu g/\mu l$. $1\mu l$ of plasmid was used for each cloning procedure.

Cloning

For some ORFs, the fragments corresponding to each ORF, previously digested and purified, were ligated in both pET22b and pGEX-KG. In a final volume of 20 μ l, a molar

ratio of 3:1 fragment/vector was ligated using 0.5 μ l of NEB T4 DNA ligase (400 units/ μ l), in the presence of the buffer supplied by the manufacturer. The reaction was incubated at room temperature for 3 hours. In some experiments, ligation was performed using the Boheringer "Rapid Ligation Kit", following the manufacturer's instructions.

In order to introduce the recombinant plasmid in a suitable strain, $100~\mu I$ *E. coli* DH5 competent cells were incubated with the ligase reaction solution for 40 minutes on ice, then at 37°C for 3 minutes, then, after adding 800 μI LB broth, again at 37°C for 20 minutes. The cells were then centrifuged at maximum speed in an Eppendorf microfuge and resuspended in approximately 200 μI of the supernatant. The suspension was then plated on LB ampicillin (100 mg/ml).

The screening of the recombinant clones was performed by growing 5 randomly-chosen colonies overnight at 37 °C in either 2 ml (pGEX or pTC clones) or 5ml (pET clones) LB broth + 100 µg/ml ampicillin. The cells were then pelletted and the DNA extracted using the Qiagen QIAprep Spin Miniprep Kit, following the manufacturer's instructions, to a final volume of 30 µl. 5 µl of each individual miniprep (approximately 1g) were digested with either NdeI/XhoI or BamHI/XhoI and the whole digestion loaded onto a 1-1.5% agarose gel (depending on the expected insert size), in parallel with the molecular weight marker (1Kb DNA Ladder, GIBCO). The screening of the positive clones was made on the base of the correct insert size.

For other ORFs, the fragments corresponding to each ORF, previously digested and purified, were ligated into both pET21b+ and pGEX-KG. A molar ratio of of 3:1 fragment/vector was used in a final volume of 20µl, that included 0.5µl T4 DNA ligase (400 units/µl, NEB) and ligation buffer supplied by the manufacturer. The reaction was performed at room temperature for 3 hours. In some experiments, ligation was performed using the Boheringer "Rapid Ligation Kit" and the manufacturer's protocol.

Recombinant plasmid was transformed into 100µl of competent *E. coli* DH5 or HB101 by incubating the ligase reaction solution and bacteria for 40 minutes on ice then at 37°C for 3 minutes. This was followed by the addition of 800µl LB broth and incubation at 37°C for 20 minutes. The cells were centrifuged at maximum speed in an Eppendorf microfuge, resuspended in approximately 200µl of the supernatant and plated onto LB ampicillin (100mg/ml) agar.

Screening for recombinant clones was performed by growing 5 randomly selected colonies overnight at 37°C in either 2.0ml (pGEX-KG clones) or 5.0ml (pET clones) LB broth + 100µg/ml ampicillin. Cells were pelleted and plasmid DNA extracted using the Qiagen QIAprep Spin Miniprep Kit, following the manufacturer's instructions. Approximately 1µg of each individual miniprep was digested with the appropriate restriction enzymes and the digest loaded onto a 1-1.5% agarose gel (depending on the expected insert size), in parallel with the molecular weight marker (1kb DNA Ladder, GIBCO). Positive clones were selected on the basis of the size of insert.

ORFs were cloned into PGEX-His, by doubly-digesting the PCR product and ligating into similarly digested vector. After cloning, recombinant plasmids were transformed into the *E.coli* host W3110. Individual clones were grown overnight at 37°C in LB broth with 50µg/ml ampicillin.

Certain ORFs may be cloned into the pGEX-HIS vector using *EcoRI-PstI* cloning sites, or *EcoRI-SalI*, or *SalI-PstI*. After cloning, the recombinant plasmids may be introduced in the *E*.coli host W3110.

Expression

Each ORF cloned into the expression vector may then be transformed into the strain suitable for expression of the recombinant protein product. 1 μl of each construct was used to transform 30 μl of *E.coli* BL21 (pGEX vector), *E.coli* TOP 10 (pTRC vector) or *E.coli* BL21-DE3 (pET vector), as described above. In the case of the pGEX-His vector, the same *E.coli* strain (W3110) was used for initial cloning and expression. Single recombinant colonies were inoculated into 2ml LB+Amp (100 μg/ml), incubated at 37°C overnight, then diluted 1:30 in 20 ml of LB+Amp (100 μg/ml) in 100 ml flasks, making sure that the OD₆₀₀ ranged between 0.1 and 0.15. The flasks were incubated at 30°C into gyratory water bath shakers until OD indicated exponential growth suitable for induction of expression (0.4-0.8 OD for pET and pTRC vectors; 0.8-1 OD for pGEX and pGEX-His vectors). For the pET, pTRC and pGEX-His vectors, the protein expression was induced by addiction of 1mM IPTG, whereas in the case of pGEX system the final concentration of IPTG was 0.2 mM. After 3 hours incubation at 30°C, the final concentration of the sample was checked by OD. In order to check expression, 1ml of each sample was removed, centrifuged in a microfuge, the pellet

resuspended in PBS, and analysed by 12% SDS-PAGE with Coomassie Blue staining. The whole sample was centrifuged at 6000g and the pellet resuspended in PBS for further use.

GST-fusion proteins large-scale purification.

For some ORFs, a single colony was grown overnight at 37°C on LB+Amp agar plate. The bacteria were inoculated into 20 ml of LB+Amp liquid colture in a water bath shaker and grown overnight. Bacteria were diluted 1:30 into 600 ml of fresh medium and allowed to grow at the optimal temperature (20-37°C) to OD₅₅₀ 0.8-1. Protein expression was induced with 0.2mM IPTG followed by three hours incubation. The culture was centrifuged at 8000 rpm at 4°C. The supernatant was discarded and the bacterial pellet was resuspended in 7.5 ml cold PBS. The cells were disrupted by sonication on ice for 30 sec at 40W using a Branson sonifier B-15, frozen and thawed two times and centrifuged again. The supernatant was collected and mixed with 150µl Glutatione-Sepharose 4B resin (Pharmacia) (previously washed with PBS) and incubated at room temperature for 30 minutes. The sample was centrifuged at 700g for 5 minutes at 4C. The resin was washed twice with 10 ml cold PBS for 10 minutes, resuspended in 1ml cold PBS, and loaded on a disposable column. The resin was washed twice with 2ml cold PBS until the flow-through reached OD₂₈₀ of 0.02-0.06. The GST-fusion protein was eluted by addition of 700µl cold Glutathione elution buffer 10mM reduced glutathione, 50mM Tris-HCl) and fractions collected until the OD_{280} was 0.1. $21\mu l$ of each fraction were loaded on a 12% SDS gel using either Biorad SDS-PAGE Molecular weight standard broad range (M1) (200, 116.25, 97.4, 66.2, 45, 31, 21.5, 14.4, 6.5 kDa) or Amersham Rainbow Marker (M") (220, 66, 46, 30, 21.5, 14.3 kDa) as standards. As the MW of GST is 26kDa, this value must be added to the MW of each GST-fusion protein.

For other ORFs, for each clone to be purified as a GST-fusion, a single colony was streaked out and grown overnight at 37°C on a LB/Amp. (100µg/ml) agar plate. An isolated colony from this plate was inoculated into 20ml of LB/Amp (100 µg/ml) liquid medium and grown overnight at 37°C with shaking. The overnight culture was diluted 1:30 into 600ml LB/Amp (100µg/ml) liquid medium and allowed to grow at the optimal temperature (20-37°C) until the OD_{550nm} reached 0.6-0.8. Recombinant protein expression was induced by addition of IPTG (final concentration 0.2mM) and the culture incubated for a further 3 hours. Bacteria were harvested by centrifugation at 8000xg for 15 min at 4°C.

;* ~

Line Room acre

The bacterial pellet was resuspended in 7.5ml cold PBS. Cells were disrupted by sonication on ice four times for 30 sec at 40W using a Branson sonifier 450 and centrifuged at 13 000xg for 30 min at 4°C. The supernatant was collected and mixed with 150µl Glutatione-Sepharose 4B resin (Pharmacia), previously equilibrated with PBS, and incubated at room temperature with gentle agitation for 30 min. The batch-wise preparation was centrifuged at 700xg for 5 min at 4°C and the supernatant discarded. The resin was washed twice (batchwise) with 10ml cold PBS for 10 min, resuspended in 1ml cold PBS, and loaded onto a disposable column. The resin continued to be washed with cold PBS, until the OD_{280nm} of the flow-through reached 0.02-0.01. The GST-fusion protein was eluted by addition of $700\mu l$ cold glutathione elution buffer (10mM reduced glutathione, 50mM Tris-HCl pH 8.0) and fractions collected, until the OD_{280nm} of the eluate indicated all the recombinant protein was obtained. 20µl aliquots of each elution fraction were analyzed by SDS-PAGE using a 12% gel. The molecular mass of the purified proteins was determined using either the Bio-Rad broad range molecular weight standard (M1) (200, 116, 97.4, 66.2, 45.0, 31.0, 21.5, 14.4, 6.5 kDa) or the Amersham Rainbow Marker (M2) (220, 66.2, 46.0, 30.0, 21.5, 14.3 kDa). The molecular weights of GST-fusion proteins are a combination of the 26 kDa GST protein and its fusion partner. Protein concentrations were estimated using the Bradford assay.

His-fusion soluble proteins large-scale purification.

For some ORFs, a single colony was grown overnight at 37°C on a LB + Amp agar ", plate. The bacteria were inoculated into 20ml of LB+Amp liquid culture and incubated overnight in a water bath shaker. Bacteria were diluted 1:30 into 600ml fresh medium and allowed to grow at the optimal temperature (20-37°C) to OD₅₅₀ 0.6-0.8. Protein expression was induced by addition of 1 mM IPTG and the culture further incubated for three hours. The culture was centrifuged at 8000 rpm at 4°C, the supernatant was discarded and the bacterial pellet was resuspended in 7.5ml cold 10mM imidazole buffer (300 mM NaCl, 50 mM phosphate buffer, 10 mM imidazole, pH 8). The cells were disrupted by sonication on ice for 30 sec at 40W using a Branson sonifier B-15, frozen and thawed two times and centrifuged again. The supernatant was collected and mixed with 150µl Ni²⁺-resin (Pharmacia) (previously washed with 10mM imidazole buffer) and incubated at room temperature with gentle agitation for 30 minutes. The sample was centrifuged at 700g for 5 minutes at 4°C. The resin was washed twice with 10 ml cold 10mM imidazole buffer for 10 minutes,

BNSDOCID: <WO__9957280A2_J_>

resuspended in 1ml cold 10mM imidazole buffer and loaded on a disposable column. The resin was washed at 4°C with 2ml cold 10mM imidazole buffer until the flow-through reached the $O.D_{280}$ of 0.02-0.06. The resin was washed with 2ml cold 20mM imidazole buffer (300 mM NaCl, 50 mM phosphate buffer, 20 mM imidazole, pH 8) until the flow-through reached the $O.D_{280}$ of 0.02-0.06. The His-fusion protein was eluted by addition of 700µl cold 250mM imidazole buffer (300 mM NaCl, 50 mM phosphate buffer, 250 mM imidazole, pH 8) and fractions collected until the $O.D_{280}$ was 0.1. $21\mu l$ of each fraction were loaded on a 12% SDS gel.

His-fusion insoluble proteins large-scale purification.

A single colony was grown overnight at 37 °C on a LB + Amp agar plate. The bacteria were inoculated into 20 ml of LB+Amp liquid culture in a water bath shaker and grown overnight. Bacteria were diluted 1:30 into 600ml fresh medium and let to grow at the optimal temperature (37°C) to O.D₅₅₀ 0.6-0.8. Protein expression was induced by addition of 1 mM IPTG and the culture further incubated for three hours. The culture was centrifuged at 8000rpm at 4°C. The supernatant was discarded and the bacterial pellet was resuspended in 7.5 ml buffer B (urea 8M, 10mM Tris-HCl, 100mM phosphate buffer, pH 8.8). The cells were disrupted by sonication on ice for 30 sec at 40W using a Branson sonifier B-15, frozen and thawed twice and centrifuged again. The supernatant was stored at -20°C, while the pellets were resuspended in 2 ml guanidine buffer (6M guanidine hydrochloride, 100mM phosphate buffer, 10 mM Tris-HCl, pH 7.5) and treated in a homogenizer for 10 cycles. The product was centrifuged at 13000 rpm for 40 minutes. The supernatant was mixed with 150 μ l Ni2+-resin (Pharmacia) (previously washed with buffer B) and incubated at room temperature with gentle agitation for 30 minutes. The sample was centrifuged at 700 g for 5 minutes at 4°C. The resin was washed twice with 10 ml buffer B for 10 minutes, resuspended in 1ml buffer B, and loaded on a disposable column. The resin was washed at room temperature with 2ml buffer B until the flow-through reached the OD_{280} of 0.02-0.06. The resin was washed with 2ml buffer C (urea 8M, 10mM Tris-HCl, 100mM phosphate buffer, pH 6.3) until the flow-through reached the O.D₂₈₀ of 0.02-0.06. The His-fusion protein was eluted by addition of 700µl elution buffer (urea 8M, 10mM Tris-HCl, 100mM phosphate buffer, pH 4.5) and fractions collected until the OD_{280} was 0.1. $21\mu l$ of each fraction were loaded on a 12% SDS gel.

A STATE OF THE STREET

Purification of His-fusion proteins.

For each clone to be purified as a His-fusion, a single colony was streaked out and grown overnight at 37°C on a LB/Amp (100 μ g/ml) agar plate. An isolated colony from this plate was inoculated into 20ml of LB/Amp (100 μ g/ml) liquid medium and grown overnight at 37°C with shaking. The overnight culture was diluted 1:30 into 600ml LB/Amp (100 μ g/ml) liquid medium and allowed to grow at the optimal temperature (20-37°C) until the OD_{550nm} reached 0.6-0.8. Expression of recombinant protein was induced by addition of IPTG (final concentration 1.0mM) and the culture incubated for a further 3 hours. Bacteria were harvested by centrifugation at 8000xg for 15 min at 4°C.

The bacterial pellet was resuspended in 7.5ml of either (i) cold buffer A (300mM NaCl, 50mM phosphate buffer, 10mM imidazole, pH 8.0) for soluble proteins or (ii) buffer B (8M urea, 10mM Tris-HCl, 100mM phosphate buffer, pH 8.8) for insoluble proteins. Cells were disrupted by sonication on ice four times for 30 sec at 40W using a Branson sonifier 450 and centrifuged at 13 000xg for 30 min at 4°C. For insoluble proteins, pellets were resuspended in 2.0 ml buffer C (6M guanidine hydrochloride, 100mM phosphate buffer, 10mM Tris-HCl, pH 7.5) and treated with a Dounce homogenizer for 10 cycles. The homogenate was centrifuged at 13 000xg for 40 min and the supernatant retained.

Supernatants for both soluble and insoluble preparations were mixed with 150µl Ni²⁺-resin (previously equilibrated with either buffer A or buffer B, as appropriate) and incubated at room temperature with gentle agitation for 30 min. The resin was Chelating Sepharose Fast Flow (Pharmacia), prepared according to manufacturers protocol. The batch-wise preparation was centrifuged at 700xg for 5 min at 4°C and the supernatant discarded. The resin was washed twice (batch-wise) with 10ml buffer A or B for 10 min, resuspended in 1.0 ml buffer A or B and loaded onto a disposable column. The resin continued to be washed with either (i) buffer A at 4°C or (ii) buffer B at room temperature, until the OD_{280nm} of the flow-through reached 0.02-0.01. The resin was further washed with either (i) cold buffer C (300mM NaCl, 50mM phosphate buffer, pH 6.3) until the the OD_{280nm} of the flow-through reached 0.02-0.01. The His-fusion protein was eluted by addition of 700µl of either (i) cold elution buffer A (300mM NaCl, 50mM phosphate buffer, 250mM imidazole, pH 8.0) or (ii) elution buffer B (8 M urea, 10mM Tris-HCl, 100mM phosphate buffer, pH 4.5) and fractions

BNSDOCID: <WO___9957280A2_i_>

collected until the $O.D_{280nm}$ indicated all the recombinant protein was obtained. $20\mu l$ aliquots of each elution fraction were analyzed by SDS-PAGE using a 12% gel. Protein concentrations were estimated using the Bradford assay.

His-fusion proteins renaturation

In the cases where denaturation was required to solubilize proteins, a renaturation step was employed prior to immunization. Glycerol was added to the denatured fractions obtained above to give a final concentration of 10%(v/v). The proteins were diluted to 200µg/ml using dialysis buffer I (10% (v/v) glycerol, 0.5M arginine, 50mM phosphate buffer, 5.0mM reduced glutathione, 0.5mM oxidised glutathione, 2.0M urea, pH 8.8) and dialysed against the same buffer for 12-14 hours at 4°C. Further dialysis was performed with buffer II (10% (v/v) glycerol, 0.5M arginine, 50mM phosphate buffer, 5.0mM reduced glutathione, 0.5mM oxidised glutathione, pH 8.8) for 12-14 hours at 4°C.

Alternatively, 10% glycerol was added to the denatured proteins. The proteins were then diluted to 20µg/ml using dialysis buffer I (10% glycerol, 0.5M arginine, 50mM phosphate buffer, 5mM reduced glutathione, 0.5mM oxidised glutathione, 2M urea, pH 8.8) and dialysed against the same buffer at 4°C for 12-14 hours. The protein was further dialysed against dialysis buffer II (10% glycerol, 0.5M arginine, 50mM phosphate buffer, 5mM reduced glutathione, 0.5mM oxidised glutathione, pH 8.8) for 12-14 hours at 4°C.

Protein concentration was evaluated using the formula:

Protein (mg/ml) =
$$(1.55 \times OD_{280}) - (0.76 \times OD_{260})$$

Purification of proteins

To analyse the solubility, pellets obtained from 3.0ml cultures were resuspended in 500µl buffer M1 (PBS pH 7.2). 25µl of lysozyme (10mg/ml) was added and the bacteria incubated for 15 min at 4°C. Cells were disrupted by sonication on ice four times for 30 sec at 40W using a Branson sonifier 450 and centrifuged at 13 000xg for 30 min at 4°C. The supernatant was collected and the pellet resuspended in buffer M2 [8M urea, 0.5M NaCl, 20mM imidazole and 0.1M NaH₂ PO₄] and incubated for 3 to 4 hours at 4°C. After centrifugation, the supernatant was collected and the pellet resuspended in buffer M3 [6M guanidinium-HCl, 0.5M NaCl, 20mM imidazole and 0.1M NaH₂PO₄] overnight at 4°C. The

supernatants from all steps were analysed by SDS-PAGE. Some proteins were found to be soluble in PBS, others need urea or guanidium-HCl for solubilization.

For preparative scale purifications, 500ml cultures were induced and fusion proteins solubilized in either buffer M1, M2 or M3 using the procedure described above. Crude extracts were loaded onto a Ni-NTA superflow column (Quiagen) equilibrated with buffer M1, M2 or M3 depending on the solubilization buffer employed. Unbound material was eluted by washing the column with the same buffer. The recombinant fusion protein was eluted with the corresponding buffer containing 500mM imidazole then dialysed against the same buffer in the absence of imidazole.

Mice immunisations

20μg of each purified protein are used to immunise mice intraperitoneally. In the case of some ORFs, Balb-C mice were immunised with Al(OH)₃ as adjuvant on days 1, 21 and 42, and immune response was monitored in samples taken on day 56. For other ORFs, CD1 mice could be immunised using the same protocol. For ORFs 25 and 40, CD1 mice were immunised using Freund's adjuvant, and the same immunisation protocol was used, except that the immune response was measured on day 42, rather than 56. Similarly, for still other ORFs, CD1 mice were immunised with Freund's adjuvant, but the immune response was measured on day 49. Alternatively, 20μg of each purified protein was mixed with Freund's adjuvant and used to immunise CD1 mice intraperitoneally. For many of the proteins, the immunization was performed on days 1, 21 and 35, and immune response was monitored in samples taken on days 34 and 49. For some proteins, the third immunization was performed on day 28, rather than 35, and the immune response was measured on days 20 and 42, rather than 34 and 49.

ELISA assay (sera analysis)

The acapsulated MenB M7 strain was plated on chocolate agar plates and incubated overnight at 37°C. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into 7ml of Mueller-Hinton Broth (Difco) containing 0.25% Glucose. Bacterial growth was monitored every 30 minutes by following OD₆₂₀. The bacteria were let to grow until the OD reached the value of 0.3-0.4. The culture was centrifuged for 10 minutes at 10000 rpm. The supernatant was discarded and bacteria were washed once with PBS, resuspended in PBS containing 0.025% formaldehyde, and incubated

for 2 hours at room temperature and then overnight at 4°C with stirring. 100 μ l bacterial cells were added to each well of a 96 well Greiner plate and incubated overnight at 4°C. The wells were then washed three times with PBT washing buffer (0.1% Tween-20 in PBS). 200 μ l of saturation buffer (2.7% Polyvinylpyrrolidone 10 in water) was added to each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT. 200 μ l of diluted sera (Dilution buffer: 1% BSA, 0.1% Tween-20, 0.1% NaN3 in PBS) were added to each well and the plates incubated for 90 minutes at 37°C. Wells were washed three times with PBT. 100 μ l of HRP-conjugated rabbit anti-mouse (Dako) serum diluted 1:2000 in dilution buffer were added to each well and the plates were incubated for 90 minutes at 37°C. Wells were washed three times with PBT buffer. 100 μ l of substrate buffer for HRP (25 ml of citrate buffer pH5, 10 mg of O-phenildiamine and 10 μ l of H₂O) were added to each well and the plates were left at room temperature for 20 minutes. 100 μ l H₂SO₄ was added to each well and OD₄₉₀ was followed. The ELISA was considered positive when OD490 was 2.5 times the respective pre-immune sera.

Alternatively, The acapsulated MenB M7 strain was plated on chocolate agar plates and incubated overnight at 37°C. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into Mueller-Hinton Broth (Difco) containing 0.25% Glucose. Bacterial growth was monitored every 30 minutes by following OD₆₂₀. The bacteria were let to grow until the OD reached the value of 0.3-0.4. The culture was centrifuged for 10 minutes at 10 000rpm. The supernatant was discarded and bacteria were washed once with PBS, resuspended in PBS containing 0.025% formaldehyde, and incubated for 1 hour at $37^{\circ}\mathrm{C}$ and then overnight at 4°C with stirring. 100µl bacterial cells were added to each well of a 96 well Greiner plate and incubated overnight at 4°C. The wells were then washed three times with PBT washing buffer (0.1% Tween-20 in PBS). $200\mu l$ of saturation buffer (2.7% Polyvinylpyrrolidone 10 in water) was added to each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT. $200\mu l$ of diluted sera (Dilution buffer: 1%BSA, 0.1% Tween-20, 0.1% NaN3 in PBS) were added to each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT. 100µl of HRP-conjugated rabbit anti-mouse (Dako) serum diluted 1:2000 in dilution buffer were added to each well and the plates were incubated for 90 minutes at 37°C. Wells were washed three times with PBT buffer. 100µl of substrate buffer for HRP (25ml of citrate buffer pH5, 10mg of O-

phenildiamine and $10\mu l$ of H_2O_2) were added to each well and the plates were left at room temperature for 20 minutes. $100\mu l$ of 12.5% H_2SO_4 was added to each well and OD_{490} was followed. The ELISA titers were calculated abitrarely as the dilution of sera which gave an OD_{490} value of 0.4 above the level of preimmune sera. The ELISA was considered positive when the dilution of sera with OD_{490} of 0.4 was higher than 1:400.

FACScan bacteria Binding Assay procedure.

The acapsulated MenB M7 strain was plated on chocolate agar plates and incubated overnight at 37°C. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into 4 tubes containing 8ml each Mueller-Hinton Broth (Difco) containing 0.25% glucose. Bacterial growth was monitored every 30 minutes by following OD_{620} . The bacteria were let to grow until the OD reached the value of 0.35-0.5. The culture was centrifuged for 10 minutes at 4000rpm. The supernatant was discarded and the pellet was resuspended in blocking buffer (1% BSA in PBS, 0.4% NaN₃) and centrifuged for 5 minutes at 4000rpm. Cells were resuspended in blocking buffer to reach OD_{620} of 0.07. 100 μl bacterial cells were added to each well of a Costar 96 well plate. 100µl of diluted (1:100, 1:200, 1:400) sera (in blocking buffer) were added to each well and plates incubated for 2 hours at 4°C. Cells were centrifuged for 5 minutes at 4000rpm, the supernatant aspirated and cells washed by addition of 200µl/well of blocking buffer in each well. 100µl of R-Phicoerytrin conjugated F(ab)₂ goat anti-mouse, diluted 1:100, was added to each well and plates incubated for 1 hour at 4°C. Cells were spun down by centrifugation at 4000rpm for 5 minutes and washed by addition of $200\mu l$ /well of blocking buffer. The supernatant was aspirated and cells resuspended in 200µl/well of PBS, 0.25% formaldehyde. Samples were transferred to FACScan tubes and read. The condition for FACScan (Laser Power 15mW) setting were: FL2 on; FSC-H threshold:92; FSC PMT Voltage: E 01; SSC PMT: 474; Amp. Gains 6.1; FL-2 PMT: 586; compensation values: 0.

OMV preparations

Bacteria were grown overnight on 5 GC plates, harvested with a loop and resuspended in 10 ml 20mM Tris-HCl. Heat inactivation was performed at 56°C for 30 minutes and the bacteria disrupted by sonication for 10' on ice (50% duty cycle, 50% output). Unbroken cells were removed by centrifugation at 5000g for 10 minutes and the total cell envelope

fraction recovered by centrifugation at 50000g at 4°C for 75 minutes. To extract cytoplasmic membrane proteins from the crude outer membranes, the whole fraction was resuspended in 2% sarkosyl (Sigma) and incubated at room temperature for 20 minutes. The suspension was centrifuged at 10000g for 10 minutes to remove aggregates, and the supernatant further ultracentrifuged at 50000g for 75 minutes to pellet the outer membranes. The outer membranes were resuspended in 10mM Tris-HCl, pH8 and the protein concentration measured by the Bio-Rad Protein assay, using BSA as a standard.

Whole Extracts preparation

Bacteria were grown overnight on a GC plate, harvested with a loop and resuspended in 1ml of 20mM Tris-HCl. Heat inactivation was performed at 56°C for 30' minutes.

Western blotting

Purified proteins (500ng/lane), outer membrane vesicles (5μg) and total cell extracts (25μg) derived from MenB strain 2996 were loaded onto a 12% SDS-polyacrylamide gel and transferred to a nitrocellulose membrane. The transfer was performed for 2 hours at 150mA at 4°C, using transfer buffer (0.3 % Tris base, 1.44 % glycine, 20% (v/v) methanol). The membrane was saturated by overnight incubation at 4°C in saturation buffer (10% skimmed milk, 0.1% Triton X100 in PBS). The membrane was washed twice with washing buffer (3% skimmed milk, 0.1% Triton X100 in PBS) and incubated for 2 hours at 37°C with mice sera diluted 1:200 in washing buffer. The membrane was washed twice and incubated for 90 minutes with a 1:2000 dilution of horseradish peroxidase labelled anti-mouse Ig. The membrane was washed twice with 0.1% Triton X100 in PBS and developed with the Opti-4CN Substrate Kit (Bio-Rad). The reaction was stopped by adding water.

Bactericidal assay

MC58 and 2996 strains were grown overnight at 37° C on chocolate agar plates. 5-7 colonies were collected and used to inoculate 7ml Mueller-Hinton broth. The suspension was incubated at 37° C on a nutator and let to grow until OD₆₂₀ was in between 0.5-0.8. The culture was aliquoted into sterile 1.5ml Eppendorf tubes and centrifuged for 20 minutes at maximum speed in a microfuge. The pellet was washed once in Gey's buffer (Gibco) and resuspended in the same buffer to an OD₆₂₀ of 0.5, diluted 1:20000 in Gey's buffer and stored at 25° C.

32

فالمنتاب والرواف فروران

50µl of Gey's buffer/1% BSA was added to each well of a 96-well tissue culture plate. 25µl of diluted (1:100) mice sera (dilution buffer: Gey's buffer/0.2% BSA) were added to each well and the plate incubated at 4°C. 25µl of the previously described bacterial suspension were added to each well. 25µl of either heat-inactivated (56°C waterbath for 30 minutes) or normal baby rabbit complement were added to each well. Immediately after the addition of the baby rabbit complement, 22µl of each sample/well were plated on Mueller-Hinton agar plates (time 0). The 96-well plate was incubated for 1 hour at 37°C with rotation and then 22µl of each sample/well were plated on Mueller-Hinton agar plates (time 1). After overnight incubation the colonies corresponding to time 0 and time 1h were counted.

Gene Variability

The *ORF4* and *919* genes were amplified by PCR on chromosomal DNA extracted from various Neisseria strains (see list of strains). The following oligonucleotides used as PCR primers were designed in the upstream and downstream regions of the genes:

orf 4.1 orf 4.3	(forward) CGAATCCGGACGGCAGGACTC (reverse) GGCAGGGAATGGCGGATTAAAG
919.1	(forward) AAAATGCCTCTCCACGGCTG or
	CTGCGCCCTGTGTTAAAATCCCCT
919.6	(reverse) CAAATAAGAAAGGAATTTTG or
	GGTATCGCAAAACTTCGCCTTAATGCG

The PCR cycling conditions were:

1 cycle	2 min. at 94°
30 cycles	30 sec. at 94°
	30 sec. at $\sim 54^{\circ}$ or $\sim 60^{\circ}$ (in according to Tm of the primers)
	40 sec. at 72°
1 cycle	7 min. at 72°

The PCR products were purified from 1 % agarose gel and sequenced using the following primers:

```
orf 4.1 (forward) CGAATCCGGACGCAGGACTC
orf 4.2 (forward) CGACCGCGCCTTTGGGACTG
orf 4.3 (reverse) GGCAGGGAATGGCGGATTAAAG
orf 4.4 (reverse) TCTTTGAGTTTGATCCAACC
```

919.1	(forward)	AAAATGCCTCTCCACGGCTG or
		CTGCGCCCTGTGTTAAAATCCCCT
919.2	(forward)	ATCCTTCCGCCTCGGCTGCG
919.3	(forward)	AAAACAGCGGCACAATCGAC
919.4	(forward)	ATAAGGGCTACCTCAAACTC
919.5	(forward)	GCGCGTGGATTATTTTTGGG
919.6	(reverse)	CAAATAAGAAAGGAATTTTG or
		GGTATCGCAAAACTTCGCCTTAATGCG
919.7	(reverse)	CCCAAGGTAATGTAGTGCCG
	(reverse)	TAAAAAAAGTTCGACAGGG
919.9	(reverse)	CCGTCCGCCTGTCGTCGCCC
919.10	(reverse)	TCGTTCCGGCGGGGTCGGGG

All documents cited herein are incorporated by reference in their entireties. The following Examples are presented to illustrate, not limit, the invention

EXAMPLE 1

Using the above-described procedures, the following oligonucleotide primers were employed in the polymerase chain reaction (PCR) assay in order to clone the ORFs as indicated:

Table 1: Oligonucleotides used for PCR for Examples 2-10

ORF	Primer	Sequence	[5-4:4:
279	Forward	CGCGGATCCCATATG-TTGCCTGCAATCACGATT	Restriction sites
		<seq 3021="" id=""></seq>	BamHI-Ndel
	Reverse		
	Reverse	CCCGCTCGAG-TTTAGAAGCGGGCGGCAA <seq< th=""><th>Xhol</th></seq<>	Xhol
İ		ID 3022>	
540	 		
519	Forward	CGCGGATCCCATATG-TTCAAATCCTTTGTCGTCA	BamHI-Ndel
		<seq 3023="" id=""></seq>	Dami II-14061
	Reverse	CCCGCTCGAG-TTTGGCGGTTTTGCTGC <seq id<="" th=""><th>Vhal</th></seq>	Vhal
		3024>	Xhol
			I
576	Forward	CGCGGATCCCATATG-GCCGCCCCCGCATCT	
		<seq 3025="" id=""></seq>	BamHI-Ndel
	Reverse	CCCGCTCGAG-ATTTACTTTTTTGATGTCGAC	
		SEQ ID 3026>	Xhol
		10EQ 10 30202	
919	Forward	CCCCCATCCCATATO	
0.0	1 Ol Walu	CGCGGATCCCATATG-TGCCAAAGCAAGAGCATC	BamHI-Ndel
	Davis	<seq 3027="" id=""></seq>	
	Reverse	CCCGCTCGAG-CGGGCGGTATTCGGG <seq id<="" th=""><th>Xhol</th></seq>	Xhol
		3028>	Alloi
444	_		
121	Forward	CGCGGATCCCATATG-GAAACACAGCTTTACAT	BamHI-Ndel
		<seq 3029="" id=""></seq>	Pallitti-Mael

	D	10000070010	
	Reverse	CCCGCTCGAG-ATAATAATATCCCGCGCCC <seq 3030="" id=""></seq>	Xhol
128	Forward	CGCGGATCCCATATG-ACTGACAACGCACT <seq 3031="" id=""></seq>	BamHI-Ndel
	Reverse	CCCG <u>CTCGAG</u> -GACCGCGTTGTCGAAA <seq 3032="" id=""></seq>	Xhol
206	Forward	CGCGGATCCCATATG-AAACACCGCCAACCGA <seq 3033="" id=""></seq>	BamHI-Ndel
	Reverse	CCCGCTCGAG-TTCTGTAAAAAAAGTATGTGC <seq 3034="" id=""></seq>	Xhol
287	Forward	CCGGAATTCTAGCTAGC-CTTTCAGCCTGCGGG <seq 3035="" id=""></seq>	EcoRi-Nhel
	Reverse	CCCGCTCGAG-ATCCTGCTCTTTTTTGCC <seq 3036="" id=""></seq>	Xhol
406	Forward	CGCGGATCCCATATG-TGCGGGACACTGACAG	BamHI-Ndel
	Reverse	CCCGCTCGAG-AGGTTGTCCTTGTCTATG <seq 3038="" id=""></seq>	Xhol

Localization of the ORFs

The following DNA and amino acid sequences are identified by titles of the following form: [g, m, or a] [#].[seq or pep], where "g" means a sequence from N. gonorrhoeae, "m" means a sequence from N. meningitidis B, and "a" means a sequence from N. meningitidis A; "#" means the number of the sequence; "seq" means a DNA sequence, and "pep" means an amino acid sequence. For example, "g001.seq" refers to an N. gonorrohoeae DNA sequence, number 1. The presence of the suffix "-1" to these sequences indicates an additional sequence found for the same ORF, thus, data for an ORF having both an unsuffixed and a suffixed sequence designation applies to both such designated sequences. Further, open reading frames are identified as ORF #, where "#" means the number of the ORF, corresponding to the number of the sequence which encodes the ORF, and the ORF designations may be suffixed with ".ng" or ".a", indicating that the ORF corresponds to a N. gonorrhoeae sequence or a N. meningitidis A sequence, respectively. The word "partial" before a sequence indicates that the sequence may be a partial or a complete ORF. Computer analysis was performed for the comparisons that follow between "g", "m", and "a" peptide sequences; and therein the "pep" suffix is implied where not expressly stated. Further, in the event of a conflict between the text immediately preceding and describing which sequences are being compared, and the

designated sequences being compared, the designated sequence controls and is the actual sequence being compared.

ORF:

contig:

279

gnm4.seq

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 3039>: m279.seq

ATAACGCGGA TTTGCGGCTG CTTGATTTCA ACGGTTTTCA GGGCTTCGGC 1 AAGTTTGTCG GCGGCGGGTT TCATCAGGCT GCAATGGGAA GGTACGGACA 51 101 CGGGCAGCGG CAGGGCGCGT TTGGCACCGG CTTCTTTGGC GGCAGCCATG 151 GCGCGTCCGA CGGCGGCGGC GTTGCCTGCA ATCACGATTT GTCCGGGTGA 201 GTTGAAGTTG ACGGCTTCGA CCACTTCGCT TTGGGCGGCT TCGGCACAAA 251 TGGCTTTAAC CTGCTCATCT TCCAAGCCGA GAATCGCCGC CATTGCGCCC 301 ACGCCTTGCG GTACGGCGGA CTGCATCAGT TCGGCGCGCA GGCGCACGAG 351 TTTGACCGCG TCGGCAAAAT TCAATGCGCC GGCGGCAACG AGTGCGGTGT 401 ATTCGCCGAG GCTGTGTCCG GCAACGGCGG CAGGCGTTTT GCCGCCCGCT 451 TCTAAATAG

This corresponds to the amino acid sequence <SEQ ID 3040; ORF 279>: m279.pep

- ITRICGCLIS TVFRASASLS AAGFIRLOWE GTDTGSGRAR LAPASLAAAM ARPTAAALPA ITICPGELKL TASTTSLWAA SAOMALTCSS SKPRIAAIAP
- TPCGTADCIS SARRRTSLTA SAKFNAPAAT SAVYSPRLCP ATAAGVLPPA
- 151 SK*

The following partial DNA sequence was identified in N.gonorrhoeae <SEQ ID 3041>: g279.seq

atgacgcgga tttgcggctg cttgatttca acggttttga gtgtttcggc 1 51 aagtttgtcg gcggcgggtt tcatcaggct gcaatgggaa ggaacggata 101 ccggcagcgg cagggcgcgt ttggctccgg cttctttggc ggcagccatg 151 gtgcgtccga cggcggcggc gttgcctgca atcacgactt gtccgggcga 201 gttgaagttg acggcttcga ccacttcgcc ctgtgcggat tcggcacaaa 251 tetgeetgae etgtteatet tecaaaceca aaatggeege cattgegeet acgcettgcg gtacggcgga ctgcatcagt tcggcgcgca ggcggacgag 351 tttgacggca tcggcaaaat ccaatgcttc ggcggcgaca agcgcggtgt attcgccgag gctgtgtccg gcaacggcgg caggcgtttt gccgcccact 401 451 tccaaatag

This corresponds to the amino acid sequence <SEQ ID 3042; ORF 279.ng>: g279.pep

- MTRICGCLIS TVLSVSASLS AAGFIRLQWE GTDTGSGRAR LAPASLAAAM VRPTAAALPA ITTCPGELKL TASTTSPCAD SAQICLTCSS SKPKMAAIAP 51 101
- TPCGTADCIS SARRRTSLTA SAKSNASAAT SAVYSPRLCP ATAAGVLPPT

ORF 279 shows 89.5% identity over a 152 aa overlap with a predicted ORF (ORF 279.ng) from N. gonorrhoeae:

```
10
                               30
                                      40
         ITRICGCLISTVFRASASLSAAGFIRLQWEGTDTGSGRARLAPASLAAAMARPTAAALPA
                                             50
m279.pep
         MTRICGCLISTVLSVSASLSAAGFIRLQWEGTDTGSGRARLAPASLAAAMVRPTAAALPA
g279
               10
                       20
                              30
                                      40
                                             50
```

```
70
                                                           80
                                                                              90
                                                                                               100
                                                                                                                  110
                                                                                                                                     120
                          ITICPGELKLTASTTSLWAASAQMALTCSSSKPRIAAIAPTPCGTADCISSARRRTSLTA
 m279.pep
                          ITTCPGELKLTASTTSPCADSAQICLTCSSSKPKMAAIAPTPCGTADCISSARRRTSLTA
 q279
                                        70
                                                           80
                                                                             90
                                                                                              100
                                                                                                                 110
                                      130
                                                         140
 m279.pep
                          SAKFNAPAATSAVYSPRLCPATAAGVLPPASKX
                          SAKSNASAATSAVYSPRLCPATAAGVLPPTSKX
 g279
                                      130
                                                         140
                                                                           150
 The following partial DNA sequence was identified in N. meningitidis <SEQ ID 3043>:
           a279.seq
                             ATGACNCNGA TTTGCGGCTG CTTGATTTCA ACGGTTTNNA GGGCTTCGGC
                        1
                      51
                             GAGTTTGTCG GCGGCGGTT TCATGAGGCT GCAATGGGAA GGTACNGACA
                             CNGGCAGCGG CAGGGCGCGT TTGGCGCCGG CTTCTTTGGC GGCAAGCATA
                             GCGCGCTCGA CGGCGGCGGC ATTGCCTGCA ATCACGACTT GTCCGGGCGA
                    201
                             GTTGAAGTTG ACGGCTTCAA CCACTTCATC CTGTGCGGAT TCGGCGCAAA
                             TTTGTTTTAC CTGTTCATCT TCCAAGCCGA GAATCGCCGC CATTGCGCCC
                    301 ACGCCTTGCG GTACGGCGGA CTGCATCAGT TCGGCGCGCA NGCGCACGAG
                    351 TTTGACCGCG TCGGCAAAAT CCAATGCGCC GGCGGCAACN AGTGCGGTGT
                    401 ATTCGCCGAN GCTGTGTCCG GCAACGGCGG CAGGCGTTTT GCCGCCCGCT
                    451 TCCGAATAG
 This corresponds to the amino acid sequence <SEQ ID 3044; ORF 279.a>:
          a279.pep
                             MTXICGCLIS TVXRASASLS AAGFMRLQWE GTDTGSGRAR LAPASLAASI
                             ARSTAAALPA ITTCPGELKL TASTTSSCAD SAQICFTCSS SKPRIAAIAP
                     51
                             TPCGTADCIS SARXRTSLTA SAKSNAPAAT SAVYSPXLCP ATAAGVLPPA
                    151 SE*
m279/a279 ORFs 279 and 279.a showed a 88.2% identity in 152 aa overlap
                                                                    20
                                                                                       30
          m279.pep
                                  ITRICGCLISTVFRASASLSAAGFIRLQWEGTDTGSGRARLAPASLAAAMARPTAAALPA
                                  pprox 1 . The property of 
          a279
                                  MTXICGCLISTVXRASASLSAAGFMRLQWEGTDTGSGRARLAPASLAASIARSTAAALPA
                                                 10
                                                                    20
                                                                                       30
                                                                                                         40
                                                                                                                            50
                                                                                                                                               60
                                                 70
                                                                    80
                                                                                       90
                                                                                                       100
                                                                                                                          110
                                  ITICPGELKLTASTTSLWAASAQMALTCSSSKPRIAAIAPTPCGTADCISSARRRTSLTA
          m279.pep
                                  ITTCPGELKLTASTTSSCADSAQICFTCSSSKPRIAAIAPTPCGTADCISSARXRTSLTA
          a279
                                                 70
                                                                   80
                                                                                       90
                                                                                                       100
                                                                                                                          110
                                                                                                                                             120
                                               130
                                                                  140
                                  SAKFNAPAATSAVYSPRLCPATAAGVLPPASKX
          m279.pep
                                  a279
                                  SAKSNAPAATSAVYSPXLCPATAAGVLPPASEX
                                               130
                                                                 140
                                                                                    150
519 and 519-1
                               gnm7.seq
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 3045>:
          m519.seq
                            (partial)
                             ..TCCGTTATCG GGCGTATGGA GTTGGACAAA ACGTTTGAAG AACGCGACGA
                      1
                                AATCAACAGT ACTGTTGTTG CGGCTTTGGA CGAGGCGGCC GGGGCTTgGG
                     51
```

GTGTGAAGGT TTTGCGTTAT GAGATTAAAG ACTTGGTTCC GCCGCAAGAA

ATCCTTCGCT CAATGCAGGC GCAAATTACT GCCGAACGCG AAAAACGCGC

CCGTATCGCC GAATCCGAAG GTCGTAAAAT CGAACAAATC AACCTTGCCA

101 151

201

```
251
                   GTGGTCAGCG CGAAGCCGAA ATCCAACAAT CCGAAGGCGA GGCTCAGGCT
                   GCGGTCAATG CGTCAAATGC CGAGAAAATC GCCCGCATCA ACCGCGCCAA
            301
                   AGGTGAAGCG GAATCCTTGC GCCTTGTTGC CGAAGCCAAT GCCGAAGCCA
            351
                   TCCGTCAAAT TGCCGCCGCC CTTCAAACCC AAGGCGGTGC GGATGCGGTC
            401
                  AATCTGAAGA TTGCGGAACA ATACGTCGCT GCGTTCAACA ATCTTGCCAA
           451
                  AGAAAGCAAT ACGCTGATTA TGCCCGCCAA TGTTGCCGAC ATCGGCAGCC
           501
           551
                  TGATTTCTGC CGGTATGAAA ATTATCGACA GCAGCAAAAC CGCCAAaTAA
 This corresponds to the amino acid sequence <SEQ ID 3046; ORF 519>:
      m519.pep
                  (partial)
                ...SVIGRMELDK TFEERDEINS TVVAALDEAA GAWGVKVLRY EIKDLVPPQE
             1
                  ILRSMQAQIT AEREKRARIA ESEGRKIEQI NLASGQREAE IQQSEGEAQA
            51
                  AVNASNAEKI ARINRAKGEA ESLRLVAEAN AEAIRQIAAA LQTQGGADAV
           101
           151
                  NLKIAEQYVA AFNNLAKESN TLIMPANVAD IGSLISAGMK IIDSSKTAK*
 The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 3047>:
                atggaatttt tcattatctt gttggcagcc gtcgccgttt tcggcttcaa
               atcetttgte gteatecece ageaggaagt ceaegttgte gaaaggeteg
            51
           101 ggcgtttcca tcgcgccctg acggccggtt tgaatatttt gattcccttt
           151 atcgaccgcg tegectaccg ccattcgctg aaagaaatcc ctttagacgt
           201 acccagccag gtctgcatca cgcgcgataa tacgcaattg actgttgacg
           251 gcatcatcta tttccaagta accgatccca aactcgcctc atacggttcg
               agcaactaca ttatggcaat tacccagett geecaaacga egetgegtte
               cgttatcggg cgtatggagt tggacaaaac gtttgaagaa cgcgacgaaa
               tcaacagtac cgtcgtctcc gccctcgatg aagccgccgg ggcttggggt
           401
          451 gtgaaagtcc tccgttacga aatcaaggat ttggttccgc cgcaagaaat
          501 cettegegea atgeaggeae aaattacege egaacgegaa aaacgegeee
          551 gtattgccga atccgaaggc cgtaaaatcg aacaaatcaa ccttgccagt
          601 ggtcagcgtg aagccgaaat ccaacaatcc gaaggcgagg ctcaggctgc
          651 ggtcaatgcg tccaatgccg agaaaatcgc ccgcatcaac cgcgccaaag
          701 gcgaagcgga atccctgcgc cttgttgccg aagccaatgc cgaagccaac
          751 cgtcaaattg ccgccgcct tcaaacccaa agcggggggg atgcggtcaa
          801 tctgaagatt gcgggacaat acgttaccgc gttcaaaaat cttgccaaag
          851 aagacaatac gcggattaag cccgccaagg ttgccgaaat cgggaaccct
               aattttcggc ggcatgaaaa attttcgcca gaagcaaaaa cggccaaata
          901
          951
This corresponds to the amino acid sequence <SEQ ID 3048; ORF 519.ng>:
     g519.pep
               MEFFIILLAA VAVFGFKSFV VIPQQEVHVV ERLGRFHRAL TAGLNILIPF
               IDRVAYRHSL KEIPLDVPSQ VCITRDNTQL TVDGIIYFQV TDPKLASYGS
           51
              SNYIMAITQL AQTTLRSVIG RMELDKTFEE RDEINSTVVS ALDEAAGAWG
          101
               VKVLRYEIKD LVPPQEILRA MQAQITAERE KRARIAESEG RKIEQINLAS
          201 GQREAEIQQS EGEAQAAVNA SNAEKIARIN RAKGEAESLR LVAEANAEAN
          251 RQIAAALQTQ SGADAVNLKI AGQYVTAFKN LAKEDNTRIK PAKVAEIGNP
          301 NFRRHEKFSP EAKTAK*
ORF 519 shows 87.5% identity over a 200 aa overlap with a predicted ORF (ORF 519.ng)
from N. gonorrhoeae:
     m519/g519
                                                        10
                                                                  20
                                                                             30
    m519.pep
                                                SVIGRMELDKTFEERDEINSTVVAALDEAA
                                                1111111111111111111111111111111111
                  YFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIGRMELDKTFEERDEINSTVVSALDEAA
    g519
                    90
                             100
                                       110
                                                 120
                                                           130
                                                                     140
                                    50
                                              60
                                                        70
                                                                  80
                  GAWGVKVLRYEIKDLVPPQEILRSMQAQITAEREKRARIAESEGRKIEQINLASGQREAE
                                                                            90
```

m519.pep

g519	 GAWGVKVLRYE 150	 KDLVPP 160	: QEILRAMQAQ: 170	 TAEREKRARI 180	 AESEGRKIEQI 190	 NLASGQREAE 200
m519.pep	100 IQQSEGEAQAA 	VNASNAE:	KIARINRAKGI	20 130 EAESLRLVAEAI	NAEAIROIAAA	LOTOGGADAV
g519	IQQSEGEAQAA 210	VNASNAE: 220	KIARINRAKGI 230	EAESLRLVAEAI 240	NAEANRQIAAA 250	LQTQSGADAV 260
m519.pep	160 NLKIAEQYVAA :	FNNLAKE	SNTLIMPANV	ADIGSL-ISAGN	90 20 MKIIDSSKTAK : :	
g519	NLKIAGQYVTA 270	FKNLAKEI 280	ONTRIKPAKVA 290	AEIGNPNFRRHI 300	EKFSPEAKTAK 310	
The following p	artial DNA seque					
1	ATGGAATTTT TCA	TTATCTT	GCTGGCAGC	GTCGTTGTT	TCGGCTTCA	A
51	ATCCTTTGTT GTC	ATCCCAC	AGCAGGAAGT	CCACGTTGT	GAAAGGCTC	G
101		CGCCCTG	ACGGCCGGTT	TGAATATTT	GATTCCCTT	T
151	ATCGACCGCG TCG	CCTACCG	CCATTCGCTC	AAAGAAATCO	CTTTAGACG	T
201	ACCCAGCCAG GTC	TGCATCA	CGCGCGACAZ	A TACGCAGCT	ACTGTTGAC	G
251	GTATCATCTA TTT	CCAAGTA	ACCGACCCC	AACTCGCCTC	CATACGGTTC	G
301	AGCAACTACA TTA	TGGCGAT	TACCCAGCTT	GCCCAAACGA	CGCTGCGTT	С
351	CGTTATCGGG CGT	ATGGAAT	TGGACAAAA	GTTTGAAGAA	A CGCGACGAA	A
401	TCAACAGCAC CGT	CGTCTCC	GCCCTCGAT	AAGCCGCCGG	AGCTTGGGG	T
451 501	GTGAAGGTTT TGC	GTTATGA	GATTAAAGAC	TTGGTTCCGC	CGCAAGAAA	T
551	CCTTCGCTCA ATG	CAGGCGC	AAATTACTGC	TGAACGCGAA	AAACGCGCC	С
601	GTATCGCCGA ATC	CCCAAGGT	CGTAAAATCG	AACAAATCAA	CCTTGCCAG	T
	GGTCAGCGCG AAG	CCGAAAT	CCAACAATCC	GAAGGCGAGG	CTCAGGCTG	С
651	GGTCAATGCG TCA	AATGCCG	AGAAAATCGC	CCGCATCAAC	CGCGCCAAA	G
701	GTGAAGCGGA ATC	CTTGCGC	CTTGTTGCCG	AAGCCAATGC	CGAAGCCAT	С
751	CGTCAAATTG CCG	CCGCCCT	TCAAACCCAA	GGCGGTGCGG	ATGCGGTCA	A
801	TCTGAAGATT GCG	GAACAAT	ACGTCGCCGC	GTTCAACAAT	CTTGCCAAA	G
851		GATTATG	CCCGCCAATG	TTGCCGACAT	CGGCAGCCT	G
901	ATTTCTGCCG GTA	TGAAAAT	TATCGACAGO	AGCAAAACCG	CCAAATAA	
This correspond	a to the emine ee	id soons	<ceo i<="" td=""><td>T 2050 OR</td><td>T 510 ·</td><td></td></ceo>	T 2050 OR	T 510 ·	
	s to the amino ac	ia seque	nce <2EQ 1	D 3030; OK	F 519.a>:	
a519.pep						
1	MEFFIILLAA VVV	FGFKSFV	VIPQQEVHVV	' ERLGRFHRAL	TAGLNILIP:	F
51	IDRVAYRHSL KEI	PLDVPSQ	VCITRDNTQI	TVDGIIYFQV	TDPKLASYG:	S
101	SNYIMAITQL AQT	TLRSVIG	RMELDKTFEE	RDEINSTVVS	ALDEAAGAW	G
151 201	VKVLRYEIKD LVP	POEILRS	MQAQITAERE	KRARIAESEG	RKIEQINLA	S
251	GQREAEIQQS EGE	AQAAVNA	SNAEKIARIN	RAKGEAESLR	LVAEANAEA	I
301	RQIAAALQTQ GGA ISAGMKIIDS SKT		AEQIVAAFNN	LAKESNTLIM	PANVADIGS:	Ĺ
m519/a519	ORFs 519 an	d 519.a	showed a 9	9.5% identi	ty in 199 a	aa overlap
m519.pep				10		30
morr.pep					TFEERDEINS	
a519	YFQVTDPKLAS	YGSSNYTM	זייייים ב. זיייים בו	TITITITE TO THE	TEEEDDETNO	
4020	90	100	110	120	130	
				120	130	140
	40	_	50 6	0 70	80	00
m519.pep	GAWGVKVLRYE					90
· · • • - •	1111111111			1111111111		TENEDUKEAE
a519	GAWGVKVLRYE	IKDLVPPC	EILRSMOAOT	TAEREKRARTA	TITLLITELLE ESEGRETEOTE	11111111111111111111111111111111111111
	150	160	170	180	190	200
						_00
	100	11	.0 12	0 130	140	150
					0	100

BNSDOCID: <WO___9957280A2_I_>

m519.pep	IQQSEGEAQA	AVNASNA	EKIARINRAKO	EAESLRLV	AEANAEAIRQI	AAALQTQGGADAV
a519			1 1 1 1 1 1 1 1 1 1 1	 EAESLRLV	 AEANAEAIRQI	
	240	220	230	240	250	260
	160		170 1	80	190	200
m519.pep	NLKIAEQYVA	AFNNLAKI	ESNTLIMPANV	ADTOSLIS	ACMETTOGGET	APV
a519	*	111111		11111111		1 + +
a319	MINIMEQIVA	AFNNLAKE	ESNTLIMPANV	ADIGSLISA	AGMKIIDSSKT	AKY
	270	280	290	300	310	

Further work revealed the DNA sequence identified in N. meningitidis <SEQ ID 3051>:

```
ATGGAATTTT TCATTATCTT GTTGGTAGCC GTCGCCGTTT TCGGTTTCAA
  1
 51 ATCCTTTGTT GTCATCCCAC AACAGGAAGT CCACGTTGTC GAAAGGCTGG
     GGCGTTTCCA TCGCGCCCTG ACGGCCGGTT TGAATATTTT GATTCCCTTT
151 ATCGACCGCG TCGCCTACCG CCATTCGCTG AAAGAAATCC CTTTAGACGT
201 ACCCAGCCAG GTCTGCATCA CGCGCGACAA TACGCAGCTG ACTGTTGACG
251 GCATCATCTA TTTCCAAGTA ACCGACCCCA AACTCGCCTC ATACGGTTCG
301 AGCAACTACA TTATGGCGAT TACCCAGCTT GCCCAAACGA CGCTGCGTTC
351 CGTTATCGGG CGTATGGAGT TGGACAAAAC GTTTGAAGAA CGCGACGAAA
     TCAACAGTAC TGTTGTTGCG GCTTTGGACG AGGCGGCCGG GGCTTGGGGT
451 GTGAAGGTTT TGCGTTATGA GATTAAAGAC TTGGTTCCGC CGCAAGAAAT
501 CCTTCGCTCA ATGCAGGCGC AAATTACTGC CGAACGCGAA AAACGCGCCC
551 GTATCGCCGA ATCCGAAGGT CGTAAAATCG AACAAATCAA CCTTGCCAGT
601 GGTCAGCGCG AAGCCGAAAT CCAACAATCC GAAGGCGAGG CTCAGGCTGC
651 GGTCAATGCG TCAAATGCCG AGAAAATCGC CCGCATCAAC CGCGCCAAAG
701 GTGAAGCGGA ATCCTTGCGC CTTGTTGCCG AAGCCAATGC CGAAGCCATC
751 CGTCAAATTG CCGCCGCCCT TCAAACCCAA GGCGGTGCGG ATGCGGTCAA
801 TCTGAAGATT GCGGAACAAT ACGTCGCTGC GTTCAACAAT CTTGCCAAAG
851 AAAGCAATAC GCTGATTATG CCCGCCAATG TTGCCGACAT CGGCAGCCTG
901 ATTTCTGCCG GTATGAAAAT TATCGACAGC AGCAAAACCG CCAAATAA
```

This corresponds to the amino acid sequence <SEQ ID 3052; ORF 519-1>:

```
1 MEFFIILLVA VAVFGFKSFV VIPQQEVHVV ERLGRFHRAL TAGLNILIPF
51 IDRVAYRHSL KEIPLDVPSQ VCITRDNTQL TVDGIIYFQV TDPKLASYGS
101 SNYIMAITQL AQTTLRSVIG RMELDKTFEE RDEINSTVVA ALDEAAGAWG
151 VKVLRYEIKD LVPPQEILRS MQAQITAERE KRARIAESEG RKIEQINLAS
201 GQREAEIQQS EGEAQAAVNA SNAEKIARIN RAKGEAESLR LVAEANAEAI
251 RQIAAALQTQ GGADAVNLKI AEQYVAAFNN LAKESNTLIM PANVADIGSL
301 ISAGMKIIDS SKTAK*
```

The following DNA sequence was identified in N. gonorrhoeae <SEQ ID 3053>:

-		4				
	1	ATGGAATTTT	TCATTATCTT	GTTGGCAGCC	GTCGCCGTTT	TCGGCTTCAA
	51	ATCCTTTGTC	GTCATCCCCC		CCACGTTGTC	
	101	GGCGTTTCCA	TCGCGCCCTG	ACGGCCGGTT		
	151	ATCGACCGCG	TCGCCTACCG		AAAGAAATCC	
	201	ACCCAGCCAG	GTCTGCATCA		TACGCAATTG	
	251	GCATCATCTA	TTTCCAAGTA	TOTOGOTITES	AACTCGCCTC	ACTGTTGACG
	301		TTATGGCAAT			
	351		CGTATGGAGT		GCCCAAACGA	
	401		CGTCGTCTCC		GTTTGAAGAA	
	451	GTGAAAGTCC	TCCCTTTT CC	GCCCTCGATG	AAGCCGCCGG	GGCTTGGGGT
	501	CCTTCCCCCA	A TOCGI I ACGA	AATCAAGGAT	TTGGTTCCGC	CGCAAGAAAT
	551	CULTUGUGUA	ATGCAGGCAC	AAATTACCGC	CGAACGCGAA	AAACGCGCCC
	601	CCTCTCCCCA	ATCCGAAGGC	CGTAAAATCG	AACAAATCAA	CCTTGCCAGT
		GGICAGCGIG	AAGCCGAAAT	CCAACAATCC	GAAGGCGAGG	CTCAGGCTGC
		GGTCAATGCG		AGAAAATCGC	CCGCATCAAC	CGCGCCAAAG
		GCGAAGCGGA		CTTGTTGCCG	AAGCCAATGC	CGAAGCCATC
		CGTCAAATTG		TCAAACCCAA		ATGCGGTCAA
	801	TCTGAAGATT	GCGGAACAAT	ACGTAGCCGC		CTTGCCAAAC
						OT TOCCAMAG

851 AAAGCAATAC GCTGATTATG CCCGCCAATG TTGCCGACAT CGGCAGCCTG ATTTCTGCCG GCATGAAAAT TATCGACAGC AGCAAAACCG CCAAATAA This corresponds to the amino acid sequence <SEQ ID 3054; ORF 519-1.ng>: g519-1.pep 1 MEFFIILLAA VAVFGFKSFV VIPQQEVHVV ERLGRFHRAL TAGLNILIPF IDRVAYRHSL KEIPLDVPSQ VCITRONTQL TVDGIIYFQV TDPKLASYGS SNYIMAITQL AQTTLRSVIG RMELDKTFEE RDEINSTVVS ALDEAAGAWG VKVLRYEIKD LVPPQEILRA MQAQITAERE KRARIAESEG RKIEQINLAS GQREAEIQQS EGEAQAAVNA SNAEKIARIN RAKGEAESLR LVAEANAEAI RQIAAALQTQ GGADAVNLKI AEQYVAAFNN LAKESNTLIM PANVADIGSL 251 301 ISAGMKIIDS SKTAK* m519-1/q519-1ORFs 519-1 and 519-1.ng showed a 99.0% identity in 315 aa overlap 10 20 30 40 MEFFIILLAAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL q519-1.pep m519-1MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL 20 30 40 70 80 90 100 110 120 KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG q519-1.pep KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG m519-180 90 100 130 140 150 160 170 180 ${\tt RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRAMQAQITAERE}$ g519-1.pep m519-1RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE 130 140 150 160 170 190 200 210 220 230 240 KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR g519-1.pep m519-1 KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR 190 200 210 220 230 250 260 270 280 LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL g519-1.pep LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL m519-1 250 260 270 280 290 310 ISAGMKIIDSSKTAKX g519-1.pep 11111111111111111 m519-1**ISAGMKIIDSSKTAKX** 310 The following DNA sequence was identified in N. meningitidis <SEQ ID 3055>: a519-1.seq ATGGAATTTT TCATTATCTT GCTGGCAGCC GTCGTTGTTT TCGGCTTCAA 1 ATCCTTTGTT GTCATCCCAC AGCAGGAAGT CCACGTTGTC GAAAGGCTCG 51 GGCGTTTCCA TCGCGCCCTG ACGGCCGGTT TGAATATTTT GATTCCCTTT 101 ATCGACCGCG TCGCCTACCG CCATTCGCTG AAAGAAATCC CTTTAGACGT 201 ACCCAGCCAG GTCTGCATCA CGCGCGACAA TACGCAGCTG ACTGTTGACG

GTATCATCTA TTTCCAAGTA ACCGACCCCA AACTCGCCTC ATACGGTTCG AGCAACTACA TTATGGCGAT TACCCAGCTT GCCCAAACGA CGCTGCGTTC

CGTTATCGGG CGTATGGAAT TGGACAAAAC GTTTGAAGAA CGCGACGAAA

BNSDOCID: <WO___9957280A2_I_>

251

301

351

....ar. ar. ar. ar. ar. ar.

aa

401	MCARCAGONG COM					
401 451	TCAACAGCAC CGTCC	STCTCC GCC	CTCGATG .	AAGCCGCCGG	AGCTTGGGGT	
501	010111111100	LIAIGA GAT	LAAAGSAC '	\mathbf{r}	~~~~~~	
551	CCTTCGCTCA ATGC	AGGCGC AAA	TACTGC '	TGAACGCGAA	AAACGCGCCC	
601	GTATCGCCGA ATCCC GGTCAGCGCG AAGCC	CAAAT CCAA	AAAATCG A	AACAAATCAA	CCTTGCCAGT	
651	GGTCAATGCG TCAAA	ATGCCG AGAZ	CAAICC (GAAGGCGAGG	CTCAGGCTGC	
701	arearesocour MICCI	. 1666 CTTT	same control of	77CCC777mcc	~~~~~~	
751	OCTONERS COGCE	GCCCT TCAL	ACCCCAA C	CCCCCCCCC	7.0000me	
801	- 0 - 0 - 1 - 1 - G C G G F	MCAAL ALGI	waren e	- או אי אי אי אי איייייי	~~~~~~~~	
851	MANGCANIAC GCIGA	MITATE CCCC	CCAATC	PTCCCCACAM	00000	
901	ATTTCTGCCG GTATG	AAAAT TATC	GACAGC A	AGCAAAACCG	CCAAATAA	
This corresponds	to the amino acid	sequence <	SEO ID	3056: ODI	2 £10 1	
apra-1.bet).					
. 1	MEFFIILLAA VVVFG	FKSFV VIPO	OEVHVV E	RIGREHRAT.	TACINITIES	
31	TOWANTEUDD VETAP	DVPSO VCTT	RONTOT. T	びかたててひをへび	MADELLE PARTE	
	OWITH THE PARTY OF TAXABLE	ROVILL RMEIL	DKTFFF D	מל זול זותים לא דים כו	77777	
101	AMATRICIAN PAREO	LILKS MOAO	TTAERE K	DADIACCCC	DESTRUCTION	
	ASTURNIT AND TREETING	AAVNA SNA⊁:	KIARINI D	カシベビカロウェゥ	T T T T T T T T T T T T T T T T T T T	
	RQIAAALQTQ GGADA' ISAGMKIIDS SKTAK	ANTRI VEOA	VAAFNN L	AKESNTLIM	PANVADIGSL	
301	ISAGMATIDS SKTAK	*				
m519-1/a51	9-1 ORFs 51	$9-1$ and 5°	19-1 a e	bound a n	9.0% identity	
overlap			1,4 3	nowed a 9	9.0% identity	y in 315 a
a519-1.pep	10	20	30	40	50	60
aJI9-I.pep		FGFKSFVVIP	QQEVHVVE	RLGRFHRALT	ACT NITT TENTENTS	
m519-1						
	MEFFIILLVAVAVE	GEKSEVVIPÇ 20	50F AHAAF1	RLGRFHRALT	AGLNILIPFIDRV	AYRHSL
	10	20	30	40	50	60 .
	70	. 80	90	100	110	100
a519-1.pep	KEIPLDVPSQVCIT	RDNTQLTVDC	SILYFQVTI	DEL VERGO	7377373	120
m519-1						
M319-1		TONI OUT A DO	SIIYFQVTI	PKLASYGSSI	YIMAITOLAOTTI	LRSVIG
	70	80	90	100	110	120
	130	140				
a519-1.pep	RMELDKTFEERDET	THU NGTWANT DE	150	160	170	180
• •	RMELDKTFEERDEI					
m519-1	RMELDKTFEERDEI	NSTVVAALDE	AAGAWGVK			
	130	140	150	160	PPQEILRSMQAQI 170	
				200	170	180
a519-1.pep	190	200	210	220	230	240
asis-i.pep	KRARIAESEGRKIE	QINLASGORE	AEIQQSEG	EAQAAVNASN		
m519-1						
	KRARIAESEGRKIE	QINLASGQRE. 200	UTT T T T O D T C	LAQAAVNASN	AEKIARINRAKGE	AESLR
	130	200	210	220	230	240
	250	260	270	280	200	
a519-1.pep	LVAEANAEAIRQIA	AALOTOGGAD	AUNIETAR	OVIII A ESINTE S	290	300
510 -						
m519-1		AALQTQGGAD	AVNLKIAE	QYVAAFNNLA:	KESNTI TMPANVA	
	250	260	270	280	290	300
	210					500
a519-1.pep	310 ISAGMKIIDSSKTAK	·v				
tob		<u> </u>				
m519-1	ISAGMKIIDSSKTAK	rx				
	310					

```
576 and 576-1 gnm22.seq
```

```
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 3057>:
```

```
m576.seq.. (partial)
          ..ATGCAGCAGG CAAGCTATGC GATGGGCGTG GACATCGGAC GCTCCCTGAA
       1
      51
            GCAAATGAAG GAACAGGGCG CGGAAATCGA TTTGAAAGTC TTTACCGAAG
            CCATGCAGGC AGTGTATGAC GGCAAAGAAA TCAAAATGAC CGAAGAGCAG
     101
            GCTCAGGAAG TCATGATGAA ATTCCTTCAG GAACAACAGG CTAAAGCCGT
     151
           AGAAAAACAC AAGGCGGACG CGAAGGCCAA TAAAGAAAAA GGCGAAGCCT
     201
            TTCTGAAAGA AAATGCCGCC AAAGACGGCG TGAAGACCAC TGCTTCCGGC
     301
            CTGCAATACA AAATCACCAA ACAGGGCGAA GGCAAACAGC CGACCAAAGA
     351
            CGACATCGTT ACCGTGGAAT ACGAAGGCCG CCTGATTGAC GGTACGGTAT
     401
            TCGACAGCAG CAAAGCCAAC GGCGGCCCGG TCACCTTCCC TTTGAGCCAA
            GTGATTCCGG GTTGGACCGA AGGCGTACAG CTTCTGAAAG AAGGCGGCGA
     451
           AGCCACGTTC TACATCCCGT CCAACCTTGC CTACCGCGAA CAGGGTGCGG
     501
           GCGACAAAAT CGGTCCGAAC GCCACTTTGG TATTTGATGT GAAACTGGTC
     551
           AAAATCGGCG CACCCGAAAA CGCGCCCGCC AAGCAGCCGG CTCAAGTCGA
     601
            CATCAAAAA GTAAATTAA
     651
```

This corresponds to the amino acid sequence <SEQ ID 3058; ORF 576>:

```
m576.pep. (partial)

1 .MQQASYAMGV DIGRSLKQMK EQGAEIDLKV FTEAMQAVYD GKEIKMTEEQ
51 AQEVMMKFLQ EQQAKAVEKH KADAKANKEK GEAFLKENAA KDGVKTTASG
101 LQYKITKQGE GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPVTFPLSQ
151 VIPGWTEGVQ LLKEGGEATF YIPSNLAYRE QGAGDKIGPN ATLVFDVKLV
201 KIGAPENAPA KQPAQVDIKK VN*
```

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 3059>: g576.seq. (partial)

```
..atgggcgtgg acatcggacg ctccctgaaa caaatgaagg aacagggcgc
       ggaaatcgat ttgaaagtct ttaccgatgc catgcaggca gtgtatgacg
 51
101
       gcaaagaaat caaaatgacc gaagagcagg cccaggaagt gatgatgaaa
       ttcctgcagg agcagcaggc taaagccgta gaaaaacaca aggcggatgc
151
201
       gaaggccaac aaagaaaaag gcgaagcctt cctgaaggaa aatgccgccg
251
       aagacggcgt gaagaccact gcttccggtc tgcagtacaa aatcaccaaa
       cagggtgaag gcaaacagcc gacaaaagac gacatcgtta ccgtggaata
301
351
       cgaaggccgc ctgattgacg gtaccgtatt cgacagcagc aaagccaacg
       gcggcccggc caccttccct ttgagccaag tgattccggg ttggaccgaa
401
       ggcgtacggc ttctgaaaga aggcggcgaa gccacgttct acatcccgtc
451
501
       caacettgee tacegegaac agggtgeggg egaaaaaate ggteegaacg
551
       ccactttggt atttgacgtg aaactggtca aaatcggcgc acccgaaaac
601
       gcgcccgcca agcagccgga tcaagtcgac atcaaaaaag taaattaa
```

This corresponds to the amino acid sequence <SEQ ID 3060; ORF 576.ng>:

```
g576.pep.(partial)

1 .MGVDIGRSLK QMKEQGAEID LKVFTDAMQA VYDGKEIKMT EEQAQEVMMK
51 FLQEQQAKAV EKHKADAKAN KEKGEAFLKE NAAEDGVKTT ASGLQYKITK
101 QGEGKQPTKD DIVTVEYEGR LIDGTVFDSS KANGGPATFP LSQVIPGWTE
151 GVRLLKEGGE ATFYIPSNLA YREQGAGEKI GPNATLVFDV KLVKIGAPEN
201 APAKQPDQVD IKKVN*
```

Computer analysis of this amino acid sequence gave the following results: Homology with a predicted ORF from N. gonorrhoeae

```
m576/g576 ORFs 576 and 576.ng showed a 97.2% identity in 215 aa overlap

10 20 30 40 50 60

m576.pep MQQASYAMGVDIGRSLKQMKEQGAEIDLKVFTEAMQAVYDGKEIKMTEEQAQEVMMKFLQ
```

g576	MGVDIGRSLKQMKEQGAEIDLKVFTDAMQAVYDGKEIKMTEEQAQEVMMKFLQ 10 20 30 40 50
m576.pep	70 80 90 100 110 120 EQQAKAVEKHKADAKANKEKGEAFLKENAAKDGVKTTASGLQYKITKQGEGKQPTKDDIV
g576	
m576.pep	
g576	TVEYEGRLIDGTVFDSSKANGGPATFPLSQVIPGWTEGVRLLKEGGEATFYIPSNLAYRE 120 130 140 150 160 170
	190 200 210 220
m576.pep	QGAGDKIGPNATLVFDVKLVKIGAPENAPAKQPAQVDIKKVNX
g576	QGAGEKIGPNATLVFDVKLVKIGAPENAPAKQPDQVDIKKVNX 180 190 200 210
The following r	partial DNA sequence was identify 1: 12
a576.seq	partial DNA sequence was identified in N. meningitidis <seq 3061="" id="">:</seq>
1 as 70.5eq	ATGAACACCA TTTTCAAAAT CAGCGCACTG ACCCTTTCCG CCGCTTTGGC
51	ACTITICGCC TECEGCAAAA AAGAAGCCGC CCCCGCATCT GCATCCGAAC
101	CTGCCGCCGC TTCTTCCGCG CAGGGCGACA CCTCTTCGAT CGGCAGCACG
1 51	ATGCAGCAGG CAAGCTATGC GATGGGGCGTG GACATCGGAC GCTCCCTGAA
201	GCAAATGAAG GAACAGGGCG CGGAAATCGA TTTGAAAGTC TTTACCGAAG
251	CCATGCAGGC AGTGTATGAC GGCAAAGAAA TCAAAATCAC CCAAGAGAGA
301	GCICAGGAAG TCATGATGAA ATTCCTTCAG GAACAACACC CERARAGGGG
351	AGARAMACAC AAGGCGGACG CGAAGGCCAA TAAACAAAAA CCCCAAACAC
401	TICIGAAAGA AAATGCCGCC AAAGACGGCG TCAACACCAC MCCMMGGGGG
451	CIGCARIACA ARATCACCAA ACAGGGGGAAA GGGAAACACA GGAAGAAA
501	COACAICGII ACCGIGGAAT ACGAAGGCCG CCTCATTCAC CCTAGGCTAG
5 51	1 COACAGCAG CAAAGCCAAC GGCGGCCCGG TCACCTTCCC TTTCACACACA
601	GIGATICIGG GITGGACCGA AGGCGTACAG CTTCTCAAAC AACCGGGG
651	AGCCACGIIC IACATCCCGI CCAACCTTGC CTACCCCCAA CACCCTGGG
701	GCGACAAAAI CGGCCCGAAC GCCACTTTGG TATTTCATCT CAAACTCC
751	AAAATOGGOG CACCOGAAAA CGCGCCCGCC AAGCAGCCCC CMCAACMGCA
801	CATCAAAAAA GTAAATTAA
This correspond	s to the amino acid sequence <seq 3062;="" 576.a="" id="" orf="">:</seq>
1	MNTTEKISAI TISAALAISA GOOTTAAAA
51	MNTIFKISAL TLSAALALSA CGKKEAAPAS ASEPAAASSA QGDTSSIGST
101	MOQASYAMGV DIGRSLKOMK EQGAEIDLKV FTEAMQAVYD GKEIKMTEEQ AQEVMMKFLQ EQQAKAVEKH KADAKANKEK GEAFLKENAA KDGVKTTASG
151	LQYKITKQGE GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPVTFPLSQ
201	VILGWTEGVQ LLKEGGEATF YIPSNLAYRE QGAGDKIGPN ATLVFDVKLV
251	KIGAPENAPA KQPAQVDIKK VN*
m576/a576	ORFs 576 and 576.a showed a 99.5% identity in 222 aa overlap
m576.pep	10 20 30
	MQQASYAMGVDIGRSLKQMKEQGAEIDLKV
a576	CGKKEAAPASASEPAAASSAOGDTSSIGSTMOODSVANGUDTGDAAAASSAOGDTSSIGSTMOODSVANGUDTGDAAAASSAOGDTSSIGSTMOODSVANGUDTGDAAAASSAOGDTSSIGSTMOODSVANGUDTGDAAAASSAOGDTSSIGSTMOODSVANGUDTGDAAAASSAOGDTSSIGSTMOODSVANGUDTGDAAAAASSAOGDTSSIGSTMOODSVANGUDTGDAAAAASSAOGDTSSIGSTMOODSVANGUDTGDAAAAASAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
	CGKKEAAPASASEPAAASSAQGDTSSIGSTMQQASYAMGVDIGRSLKQMKEQGAEIDLKV 30 40 50 60 70
	30 40 50 60 70 80
_	40 50 60 70 80 90
m576.pep	FTEAMQAVYDGKEIKMTEEOAOEVMMKFLOFOOAKAVEVUVA DAVAVVEVUVA DAVAVEVU VA DAVAVEVU VA DAVAVVEVU VA DAVAVEVU VA DAVAVEVU VA DAVAVE VU VA DAVAVE VA DAVA DA
- 57.0	
a576	
	90 100 110 120 130 140

m576.pep	100 KDGVKTTASGLQYKI	110	120 PTKDDIVTVEY	130	140	150
	111111111111111			111111111	IIIIIIII	ALEETPOO
a576	KDGVKTTASGLQYKI	TKQGEGKQ	PTKDDIVTVEY	EGRLIDGTVE	DSSKANGGP	
	150	160	170	180	190	200
	1.60	1.00				
	160	170	180	190	200	210
m576.pep	VIPGWTEGVQLLKEG	GEATFYIP	SNLAYREQGAG	DKIGPNATLV	FDVKLVKIG	APENAPA
	11 11111111111	11111111	1111111111	[]]]]]	111111111	11111
a576	VILGWTEGVQLLKEG	GEATFYIP	SNLAYREOGAGI	DKIGPNATLV	FDVKLVKTG	APENAPA
	210	220	230	240	250	260
~					250	200
	220					
m576.pep	KOPAOVDIKKVNX					
• •	11111111111					
a576	KOPAOVDIKKVNX	•				
40.0	270					
	210					

Further work revealed the DNA sequence identified in N. meningitidis <SEQ ID 3063>:

```
m576-1.seq
      1 ATGAACACCA TTTTCAAAAT CAGCGCACTG ACCCTTTCCG CCGCTTTGGC
      51 ACTTTCCGCC TGCGGCAAAA AAGAAGCCGC CCCCGCATCT GCATCCGAAC
     101 CTGCCGCCGC TTCTTCCGCG CAGGGCGACA CCTCTTCGAT CGGCAGCACG
     151 ATGCAGCAGG CAAGCTATGC GATGGGCGTG GACATCGGAC GCTCCCTGAA
201 GCAAATGAAG GAACAGGGCG CGGAAATCGA TTTGAAAGTC TTTACCGAAG
     251 CCATGCAGGC AGTGTATGAC GGCAAAGAAA TCAAAATGAC CGAAGAGCAG
     301 GCTCAGGAAG TCATGATGAA ATTCCTTCAG GAACAACAGG CTAAAGCCGT
     351 AGAAAAACAC AAGGCGGACG CGAAGGCCAA TAAAGAAAAA GGCGAAGCCT
     401 TTCTGAAAGA AAATGCCGCC AAAGACGGCG TGAAGACCAC TGCTTCCGGC
          CTGCAATACA AAATCACCAA ACAGGGCGAA GGCAAACAGC CGACCAAAGA
     451
     501 CGACATCGTT ACCGTGGAAT ACGAAGGCCG CCTGATTGAC GGTACGGTAT
     551 TCGACAGCAG CAAAGCCAAC GGCGGCCCGG TCACCTTCCC TTTGAGCCAA
     601 GTGATTCCGG GTTGGACCGA AGGCGTACAG CTTCTGAAAG AAGGCGGCGA
     651 AGCCACGTTC TACATCCCGT CCAACCTTGC CTACCGCGAA CAGGGTGCGG
     701
         GCGACAAAAT CGGTCCGAAC GCCACTTTGG TATTTGATGT GAAACTGGTC
          AAAATCGGCG CACCCGAAAA CGCGCCCGCC AAGCAGCCGG CTCAAGTCGA
     751
     801 CATCAAAAA GTAAATTAA
```

This corresponds to the amino acid sequence <SEQ ID 3064; ORF 576-1>:

m576-1.pep 1 N

MNTIFKISAL	TLSAALALSA	CGKKEAAPAS	ASEPAAASSA	QGDTSSIGST
MQQASYAMGV	DIGRSLKOMK	EQGAEIDLKV	FTEAMQAVYD	GKEIKMTEEO
AQEVMMKFLQ	EQQAKAVEKH	KADAKANKEK	GEAFLKENAA	KDGVKTTASG
LQYKITKQGE	GKQPTKDDIV	TVEYEGRLID	GTVFDSSKAN	GGPVTFPLSO
VIPGWTEGVQ	LLKEGGEATF	YIPSNLAYRE	OGAGDKIGPN	ATLVFDVKLV
			-	
	MQQASYAMGV AQEVMMKFLQ LQYKITKQGE VIPGWTEGVQ	MQQASYAMGV DIGRSIKQMK AQEVMMKFLQ EQQAKAVEKH LQYKITKQGE GKQPTKDDIV VIPGWTEGVQ LLKEGGEATF	MQQASYAMGV DIGRSLKQMK EQGAEIDLKV AQEVMMKFLQ EQQAKAVEKH KADAKANKEK LQYKITKQGE GKQPTKDDIV TVEYEGRLID	VIPGWTEGVQ LLKEGGEATF YIPSNLAYRE QGAGDKIGPN

The following DNA sequence was identified in N. gonorrhoeae <SEQ ID 3065>: g576-1.seq

1	ATGAACACCA	TTTTCAAAAT	CAGCGCACTG	ACCCTTTCCG	CCGCTTTGGC
51	ACTTTCCGCC	TGCGGCAAAA	AAGAAGCCGC	CCCCGCATCT	GCATCCGAAC
101	CTGCCGCCGC	TTCTGCCGCG	CAGGGCGACA	CCTCTTCAAT	CGGCAGCACG
151	ATGCAGCAGG	CAAGCTATGC	AATGGGCGTG	GACATCGGAC	GCTCCCTGAA
201		GAACAGGGCG			
251	CCATGCAGGC	AGTGTATGAC	GGCAAAGAAA	TCAAAATGAC	CGAAGAGCAG
301	GCCCAGGAAG	TGATGATGAA	ATTCCTGCAG	GAGCAGCAGG	CTAAAGCCGT
351	AGAAAAACAC	AAGGCGGATG	CGAAGGCCAA	CAAAGAAAA	GGCGAAGCCT
401	TCCTGAAGGA	AAATGCCGCC	AAAGACGGCG	TGAAGACCAC	TGCTTCCGGT
451	CTGCAGTACA	AAATCACCAA	ACAGGGTGAA	GGCAAACAGC	CGACAAAAGA
501	CGACATCGTT	ACCGTGGAAT	ACGAAGGCCG	CCTGATTGAC	GGTACCGTAT
551	TCGACAGCAG	CAAAGCCAAC	GGCGGCCCGG	CCACCTTCCC	TTTGAGCCAA
601		GTTGGACCGA			
651	AGCCACGTTC	TACATCCCGT	CCAACCTTGC	CTACCGCGAA	CAGGGTGCGG

BNSDOCID: <WO___9957280A2_I_>

701 GCGAAAAAAT CGGTCCGAAC GCCACTTTGG TATTTGACGT GAAACTGGTC AAAATCGGCG CACCCGAAAA CGCGCCCGCC AAGCAGCCGG ATCAAGTCGA 801 САТСАААААА GTAAATTAA This corresponds to the amino acid sequence <SEQ ID 3066; ORF 576-1.ng>: g576-1.pep MNTIFKISAL TLSAALALSA CGKKEAAPAS ASEPAAASAA QGDTSSIGST 1 MQQASYAMGV DIGRSLKOMK EQGAEIDLKV FTDAMQAVYD GKEIKMTEEQ 51 AQEVMMKFLQ EQQAKAVEKH KADAKANKEK GEAFLKENAA KDGVKTTASG 101 LQYKITKQGE GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPATFPLSQ VIPGWTEGVR LLKEGGEATF YIPSNLAYRE QGAGEKIGPN ATLVFDVKLV 201 251 KIGAPENAPA KQPDQVDIKK VN* ORFs 576-1 and 576-1.ng showed a 97.8% identity in 272 aa g576-1/m576-1 overlap 10 20 30 40 MNTIFKISALTLSAALALSACGKKEAAPASASEPAAASAAQGDTSSIGSTMQQASYAMGV 50 g576-1.pep MNTIFKISALTLSAALALSACGKKEAAPASASEPAAASSAQGDTSSIGSTMQQASYAMGV m576-110 20 30 40 50 70 80 90 100 DIGRSLKOMKEOGAEIDLKVFTDAMOAVYDGKEIKMTEEQAQEVMMKFLQEQQAKAVEKH g576-1.pep DIGRSLKOMKEOGAEIDLKVFTEAMOAVYDGKEIKMTEEQAQEVMMKFLOEQQAKAVEKH m576-1 80 90 100 110 130 140 150 160 KADAKANKEKGEAFLKENAAKDGVKTTASGLQYKITKQGEGKQPTKDDIVTVEYEGRLID 170 g576-1.pep KADAKANKEKGEAFLKENAAKDGVKTTASGLQYKITKQGEGKQPTKDDIVTVEYEGRLID m576-1 130 140 150 160 170 190 200 210 220 GTVFDSSKANGGPATFPLSQVIPGWTEGVRLLKEGGEATFYIPSNLAYREQGAGEKIGPN 230 g576-1.pep វិណីពេកការការការបើកបានប្រជាជាក្រុមការបានប្រជាជាក្រុមការបានប្រជាក្រុមការបានប្រជាជាក្រុមការបានប្រជាជាក្រុមការបានប GTVFDSSKANGGPVTFPLSQVIPGWTEGVQLLKEGGEATFYIPSNLAYREQGAGDKIGPN m576-1190 200 210 220 230 240 250 260 g576-1.pep ATLVFDVKLVKIGAPENAPAKQPDQVDIKKVNX m576-1ATLVFDVKLVKIGAPENAPAKQPAQVDIKKVNX 250 260 270 The following DNA sequence was identified in N. meningitidis <SEQ ID 3067>: a576-1.seq ATGAACACCA TTTTCAAAAT CAGCGCACTG ACCCTTTCCG CCGCTTTGGC 1 ACTITCCGCC TGCGGCAAAA AAGAAGCCGC CCCCGCATCT GCATCCGAAC 51 CTGCCGCCGC TTCTTCCGCG CAGGGCGACA CCTCTTCGAT CGGCAGCACG 101 ATGCAGCAGG CAAGCTATGC GATGGGCGTG GACATCGGAC GCTCCCTGAA 151 GCAAATGAAG GAACAGGGCG CGGAAATCGA TTTGAAAGTC TTTACCGAAG 201 CCATGCAGGC AGTGTATGAC GGCAAAGAAA TCAAAATGAC CGAAGAGCAG 251 GCTCAGGAAG TCATGATGAA ATTCCTTCAG GAACAACAGG CTAAAGCCGT 301 AGAAAAACAC AAGGCGGACG CGAAGGCCAA TAAAGAAAAA GGCGAAGCCT 351 TTCTGAAAGA AAATGCCGCC AAAGACGGCG TGAAGACCAC TGCTTCCGGC 401 CTGCAATACA AAATCACCAA ACAGGGCGAA GGCAAACAGC CGACCAAAGA 451 501 CGACATCGTT ACCGTGGAAT ACGAAGGCCG CCTGATTGAC GGTACGGTAT TCGACAGCAG CAAAGCCAAC GGCGGCCCGG TCACCTTCCC TTTGAGCCAA 551 GTGATTCTGG GTTGGACCGA AGGCGTACAG CTTCTGAAAG AAGGCGGCGA 601 AGCCACGTTC TACATCCCGT CCAACCTTGC CTACCGCGAA CAGGGTGCGG

651

```
GCGACAAAAT CGGCCCGAAC GCCACTTTGG TATTTGATGT GAAACTGGTC
             AAAATCGGCG CACCCGAAAA CGCGCCCGCC AAGCAGCCGG CTCAAGTCGA
         801
             CATCAAAAA GTAAATTAA
This corresponds to the amino acid sequence <SEQ ID 3068; ORF 576-1.a>:
    a576-1.pep
          1 MNTIFKISAL TLSAALALSA CGKKEAAPAS ASEPAAASSA QGDTSSIGST
51 MQQASYAMGV DIGRSLKQMK EQGAEIDLKV FTEAMQAVYD GKEIKMTEEQ
         101 AQEVMMKFLQ EQQAKAVEKH KADAKANKEK GEAFLKENAA KDGVKTTASG
         151 LQYKITKQGE GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPVTFPLSQ
         201 VILGWTEGVQ LLKEGGEATF YIPSNLAYRE QGAGDKIGPN ATLVFDVKLV
251 KIGAPENAPA KQPAQVDIKK VN*
                     ORFs 576-1 and 576-1.a showed a 99.6% identity in 272 aa
    a576-1/m576-1
    overlap
                                20
                                         30
                                                  40
                                                                    60
    a576-1.pep
                MNTIFKISALTLSAALALSACGKKEAAPASASEPAAASSAQGDTSSIGSTMQQASYAMGV
                m576-1
                MNTIFKISALTLSAALALSACGKKEAAPASASEPAAASSAQGDTSSIGSTMQQASYAMGV
                                20
                                                  40
                                                           50
                       70
                                80
                                         90
                                                 100
                                                          110
                DIGRSLKQMKEQGAEIDLKVFTEAMQAVYDGKEIKMTEEQAQEVMMKFLQEQQAKAVEKH
    a576-1.pep
                DIGRSLKQMKEQGAEIDLKVFTEAMQAVYDGKEIKMTEEQAQEVMMKFLQEQQAKAVEKH
    m576-1
                       70
                                80
                                         90
                                                 100
                                                          110 .
                      130
                               140
                                        150
                                                 160
                                                          170
    a576-1.pep
               KADAKANKEKGEAFLKENAAKDGVKTTASGLQYKITKQGEGKQPTKDDIVTVEYEGRLID
                m576-1
               KADAKANKEKGEAFLKENAAKDGVKTTASGLQYKITKQGEGKQPTKDDIVTVEYEGRLID
                      130
                               140
                                        150
                                                 160
                                                          170
                      190
                               200
                                        210
                                                 220
                                                          230
               GTVFDSSKANGGPVTFPLSQVILGWTEGVQLLKEGGEATFYIPSNLAYREQGAGDKIGPN
    a576-1.pep
                GTVFDSSKANGGPVTFPLSQVIPGWTEGVQLLKEGGEATFYIPSNLAYREQGAGDKIGPN
    m576-1
                      190
                               200
                                        210
                                                 220
                                                          230
                                                                   240
                      250
                               260
               ATLVFDVKLVKIGAPENAPAKQPAQVDIKKVNX
               ATLVFDVKLVKIGAPENAPAKQPAQVDIKKVNX
   m576-1
                               260
```

919 gnm43.seq

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 3069>:

1	ATGAAAAAAT	ACCTATTCCG	CGCCGCCCTG	TACGGCATCG	CCGCCGCCAT
51	CCTCGCCGCC	TGCCAAAGCA	AGAGCATCCA	AACCTTTCCG	CAACCCGACA
101	CATCCGTCAT	CAACGGCCCG	GACCGGCCGG	TCGGCATCCC	CGACCCCGCC
151	GGAACGACGG	TCGGCGGCGG	CGGGGCCGTC	TATACCGTTG	TACCGCACCT
201	GTCCCTGCCC	CACTGGGCGG	CGCAGGATTT	CGCCAAAAGC	CTGCAATCCT
251	TCCGCCTCGG	CTGCGCCAAT	TTGAAAAACC	GCCAAGGCTG	GCAGGATGTG
301	TGCGCCCAAG	CCTTTCAAAC	CCCCGTCCAT	TCCTTTCAGG	CAAAACAGTT
351	TTTTGAACGC	TATTTCACGC	CGTGGCAGGT	TGCAGGCAAC	GGAAGCCTTG
401	CCGGTACGGT	TACCGGCTAT	TACGAACCGG	TGCTGAAGGG	CGACGACAGG



This corresponds to the amino acid sequence <SEQ ID 3070; ORF 919>:

```
1 MKKYLFRAAL YGIAAAILAA CQSKSIQTFP QPDTSVINGP DRPVGIPDPA
51 GTTVGGGGAV YTVVPHLSLP HWAAQDFAKS LQSFRLGCAN LKNRQGWQDV
101 CAQAFQTPVH SFQAKQFFER YFTPWQVAGN GSLAGTVTGY YEPVLKGDDR
151 RTAQARFPIY GIPDDFISVP LPAGLRSGKA LVRIRQTGKN SGTIDNTGGT
201 HTADLSRFPI TARTTAIKGR FEGSRFLPYH TRNQINGGAL DGKAPILGYA
251 EDPVELFFMH IQGSGRLKTP SGKYIRIGYA DKNEHPYVSI GRYMADKGYL
301 KLGQTSMQGI KSYMRQNPQR LAEVLGQNPS YIFFRELAGS SNDGPVGALG
351 TPLMGEYAGA VDRHYITLGA PLFVATAHPV TRKALNRLIM AQDTGSAIKG
401 AVRVDYFWGY GDEAGELAGK QKTTGYVWQL LPNGMKPEYR P*
```

The following partial DNA sequence was identified in N.gonorrhoeae <SEQ ID 3071>:

ra.sed				`
1	ATGAAAAAAC ACCTGCTCC	G CTCCGCCCTC	TACGGcatCG	CCCCCCCC
51	CCCCAAAqc	a gGAGCATCC		Chaccooner
101	CATCCGICAT CAACGGCCC	G GACCGGCCGG	CCCCCATCC	CCACCCCCCC
151	GGAACGACGG TTGCCGGCG	G CGGGGCCGTC	TATACCCTTC	TCCCCCCA
201	GTCCATGCCC CACTGGGCG	G CGCaggATTT	TGCCAAAACC	GECCGCACCT.
251	TCCGCCTCGG CTGCGCCAA	T TTGAAAAACC	'GCCAAGGCTTC	CIGCAATCCT
301	TGCGCCCAAG CCTTTCAAA	C CCCCGTGCAT	TOCCARGGCIG	GCAGGATGTG
351	TTTTGAACGC TATTTCACG	C catGGCaaat	tocarcal	CAAAGCGGTT
401	Caggtacggt TACCGGCTA	T TACGAACCGG	TGCTGAACCG	GGAAGCCTTG
451	CGGACGGAAC GGGCCCGCT	I CCCGATTTAC	GGTATTCCCC	ACCAMMENT
501	CTCCGTCCCG CTGCCTGCC	G GTTTGCGGG	CCCAAAAAAA	ACGATTTTAT
551	TCAGGCAGAC ggGGAAAAA	C AGCGGCACGA	TCCACAAAAAC	CTTGTCCGCA
601	CATACCGCCG ACCTCTCCCC	ATTCCCCATC	ACCCCCCCCCCC	CGGCGGCACG
651	CaaaGGCAGG TTTGAaggA	4 GCCGCTTCCT	CCCTTACCA	CAACGGCaat
701	AAAtcaacGG CGGCgcgcT	r Gacggraaag	CCCTTACCAC	ACGCGCAACC
751	GAAGACCCCG tcgaacttT	TTTCATCCAC	Atagaage	CggttacgcC
801	GAAAACCCcg tccggcaaat	acatececat	Acceaagger	CGGGCCGCCT
851	AACAtccgTa tgtttccatc	COACCCETATA	Toggaraegee	gacaaaaacg
901	AAGCtcgggc agACCTCGAT	GCACCCCAIA	IGGCGGACAA	AGGCTACCTC
951	TCCGCAACGC CTCGCCGAAC	TTTTGGGTCA	AAAGCCTATA	TGCGGCAAAA
1001	TCCGCGAGCT TGCCGGAAGC	CGCAATCACC	AAACCCCAGC	TATATCTTTT
1051	ACGCCACTGA TGGGGGAATA	CCCCCCCCCC	GCCCCGTCGG	CGCACTGGGC
1101	CTTGGGCGCG CCCTTATTTC	TCCCCA	ATCGACCGGC	ACTACATTAC
1151	CCCTCAACCG CCTGATTATG	GCCCACCGC	CCATCCGGTT	ACCCGCAAAG
1201	GCGGTGCGCG TGGATTATTT	TTCCCCTTT	CAGGCAGCGC	GATCAAAGGC
1251	GCGGTGCGCG TGGATTATTT	T T G G G T T A C	GGCGACGAAG	CCGGCGAACT
1301	TGCCGGCAAA CAGAAAACCA GCATGAAGCC CGAATACCGC	CCCTCT	CTGGCAGCTC	CTGCCCAACG
		CCGTGA		

This corresponds to the amino acid sequence <SEQ ID 3072; ORF 919.ng>:

.pep					_
1	MKKHLLRSAL	YGIAAAILAA	COSRSIQTFP	QPDTSVINGP	DRPAGIPDPA
51	GTTVAGGGAV	YTVVPHLSMP	HWAAQDFAKS	LOSFRLGCAN	LKNROGWODV
101	CAQAFQTPVH	SFQAKRFFER	YFTPWQVAGN	GSLAGTVTGY	YEPVLKGDGR
151	RTERARFPIY	GIPDDFISVP	LPAGLRGGKN	LVRIRQTGKN	SGTIDNAGGT
201	HTADLSRFPI	TARTTAIKGR	FEGSRFLPYH	TRNQINGGAL	DGKAPILGYA
251	EDPVELFFMH	IQGSGRLKTP	SGKYIRIGYA	DKNEHPYVSI	GRYMADKGYL
301	KLGQTSMQGI	KAYMRQNPQR	LAEVLGONPS	YIFFRELAGS	GNEGPVGALG
351	TPLMGEYAGA	IDRHYITLGA	PLFVATAHPV	TRKALNRLIM	AODTGSAIKG
401	AVRVDYFWGY	GDEAGELAGK	QKTTGYVWQL	LPNGMKPEYR	P*

ORF 919 shows 95.9 % identity over a 441 aa overlap with a predicted ORF (ORF 919.ng) from *N. gonorrhoeae:* m919/g919

	10	20	30	40	50	60
m919.pep	MKKYLFRAALYGIA	AAILAACQS	KSIQTFPQPD1	SVINGPDRP	GI PDPAGTT	VGGGGAV
			:			1.1111
g919	MKKHLLRSALYGIA	AAILAACOS	RSIOTFPOPDT	SVINGPDRP	AGT PD PAGTT	VAGGGAV
	10	20	30	40	50	60
					30	60
	70	80	90	100	110	100
m919.pep	YTVVPHLSLPHWAA				TIU	120
· F - F			IIIIIIIIIIII	(QGWQDVCAQA	TOTPVHSFQ.	AKOFFER
q919	:					:
9717	YTVVPHLSMPHWAA 70	ODEWY2DOS				AKRFFER
	70	80	90	100	110	120
	100					
	130	140	150	160	170	180
m919.pep	YFTPWQVAGNGSLA	GTVTGYYEP	/LKGDDRRTAC	ARFPIYGIPI	DFISVPLPA	GLRSGKA
		, , , , , , , , , ,		1111111		:
g919	YFTPWQVAGNGSLA	GTVTGYYEP	/LKGDGRRTER	ARFPIYGIPE	DFISVPLPA	GLRGGKN 🐍
•	130	140	150	160	170	180
	190	200	210	220	230	240
m919.pep	LVRIRQTGKNSGTI	DNTGGTHTAL	LSRFPITART	TAIKGRFEGS	RFLPYHTRNO	TNCCAL.
		11:111111		111111111	11111111	111111
g919	LVRIRQTGKNSGTI	DNAGGTHTAI	LSREPTTART	TATKGRFFGS		
	190	200	210	220	230	21NGGAD 240
			210	240	230	240
	250	260	270	280	290	
m919.pep	DGKAPILGYAEDPV				290	300
				INIGIADANE	HPIVSIGRY	
g919						
9727	DGKAPILGYAEDPV	260				
	250	260	270	280	290	300
	210	200				
m010 mam	310	320	330	340	350	360
m919.pep	KLGQTSMQGIKSYM	RONPORLAEV	LGONPSYIFF	RELAGSSNDG	PVGALGTPLN	<i>I</i> GEYAGA
	111111111111	!!!!!!!!!!	1111111111			
g919	KLGQTSMQG1KAYM	RQNPQRLAEV	LGQNPSYIFF	RELAGSGNEG	PVGALGTPLN	IGEYAGA
	310	320	330	340	350	360
	370	380	390	400	410	420
m919.pep	VDRHYITLGAPLFV	ATAHPVTRKA	LNRLIMAODT	GSAI KGAVRV	DYFWGYGDE	GEI.ACK
	:	[] [] [] [] [] [1111111111			
g9 19	IDRHYITLGAPLFV	ATAHPVTRKA	LNRLIMAODT	GSAIKGAVDV		IIIIII
	370	380	390	400	410	
					410	420

430 440 m919.pep QKTTGYVWQLLPNGMKPEYRPX g919 QKTTGYVWQLLPNGMKPEYRPX 430 440

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 3073>:

a919.seg ATGAAAAAAT ACCTATTCCG CGCCGCCCTG TGCGGCATCG CCGCCGCCAT CCTCGCCGCC TGCCAAAGCA AGAGCATCCA AACCTTTCCG CAACCCGACA 101 CATCCGTCAT CAACGGCCCG GACCGGCCGG TCGGCATCCC CGACCCCGCC 151 GGAACGACGG TCGGCGGCGG CGGGGCCGTT TATACCGTTG TGCCGCACCT GTCCCTGCCC CACTGGGCGG CGCAGGATTT CGCCAAAAGC CTGCAATCCT 251 TCCGCCTCGG CTGCGCCAAT TTGAAAAACC GCCAAGGCTG GCAGGATGTG 301 TGCGCCCAAG CCTTTCAAAC CCCCGTCCAT TCCGTTCAGG CAAAACAGTT 351 TTTTGAACGC TATTTCACGC CGTGGCAGGT TGCAGGCAAC GGAAGCCTTG 401 CCGGTACGGT TACCGGCTAT TACGAGCCGG TGCTGAAGGG CGACGACAGG CGGACGGCAC AAGCCCGCTT CCCGATTTAC GGTATTCCCG ACGATTTTAT 451 501 CTCCGTCCC CTGCCTGCCG GTTTGCGGAG CGGAAAAGCC CTTGTCCGCA 551 TCAGGCAGAC GGGAAAAAAC AGCGGCACAA TCGACAATAC CGGCGGCACA 601 CATACCGCCG ACCTCTCCCA ATTCCCCATC ACTGCGCGCA CAACGGCAAT 651 CAAAGGCAGG TTTGAAGGAA GCCGCTTCCT CCCCTACCAC ACGCGCAACC 701 AAATCAACGG CGGCGCGCTT GACGGCAAAG CCCCGATACT CGGTTACGCC GAAGACCCCG TCGAACTTTT TTTTATGCAC ATCCAAGGCT CGGGCCGTCT 751 801 GAAAACCCCG TCCGGCAAAT ACATCCGCAT CGGCTATGCC GACAAAAACG 851 AACATCCCTA CGTTTCCATC GGACGCTATA TGGCGGACAA AGGCTACCTC 901 AAGCTCGGGC AGACCTCGAT GCAGGGCATC AAAGCCTATA TGCAGCAAAA 951 CCCGCAACGC CTCGCCGAAG TTTTGGGGCA AAACCCCAGC TATATCTTTT 1001 TCCGAGAGCT TACCGGAAGC AGCAATGACG GCCCTGTCGG CGCACTGGGC 1051 ACGCCGCTGA TGGGCGAGTA CGCCGGCGCA GTCGACCGGC ACTACATTAC 1101 CTTGGGCGCG CCCTTATTTG TCGCCACCGC CCATCCGGTT ACCCGCAAAG 1151 CCCTCAACCG CCTGATTATG GCGCAGGATA CCGGCAGCGC GATTAAAGGC 1201 GCGGTGCGCG TGGATTATTT TTGGGGATAC GGCGACGAAG CCGGCGAACT TGCCGGCAAA CAGAAAACCA CGGGATATGT CTGCCCAACG GTATGAAGCC CGAATACCGC CCGTAA 1301

This corresponds to the amino acid sequence <SEQ ID 3074; ORF 919.a>: a919.pep

MKKYLFRAAL CGIAAAILAA CQSKSIQTFP QPDTSVINGP DRPVGIPDPA GTTVGGGGAV YTVVPHLSLP HWAAQDFAKS LQSFRLGCAN LKNRQGWQDV 51 CAQAFQTPVH SVQAKQFFER YFTPWQVAGN GSLAGTVTGY YEPVLKGDDR RTAQARFPIY GIPDDFISVP LPAGLRSGKA LVRIRQTGKN SGTIDNTGGT HTADLSQFPI TARTTAIKGR FEGSRFLPYH TRNQINGGAL DGKAPILGYA 251 EDPVELFFMH IQGSGRLKTP SGKYIRIGYA DKNEHPYVSI GRYMADKGYL 301 KLGQTSMQGI KAYMQQNPQR LAEVLGQNPS YIFFRELTGS SNDGPVGALG TPLMGEYAGA VDRHYITLGA PLFVATAHPV TRKALNRLIM AQDTGSAIKG 401 AVRVDYFWGY GDEAGELAGK QKTTGYVWQL LPNGMKPEYR P*

m919/a919 ORFs 919 and 919.a showed a 98.6% identity in 441 aa overlap 20

m919.pep MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVG	
a919 MKKYLFRAALCGIAAATLAACOSKSTOTEPODDTSVINODDD DVGT	1111
	~~~~
10 · 20 30	GGAV .
10 · 20 30 40 50	60
70 80 90 100 110	100
	120
	OFFER
4919 YTWYPHI SI DHWAAODEANGI OGERT GOODEANGI OGERT	
11 VI MESTE HWAAQDEAKSLQSEKLGCANLKNRQGWODVCAOA FOT PVHSVOAK	סשששר
	SEFEK
70 80 90 100 110	120

30

m919.pep	130 YFTPWQVAGNGSLA             YFTPWQVAGNGSLA 130			111111111	11111111111	1111111
m919.pep	190 LVRIRQTGKNSGTI            LVRIRQTGKNSGTI 190	111111111		11111111	1111111111	111111
m919.pep	250 DGKAPILGYAEDPV             DGKAPILGYAEDPV 250	1111111111		11111111		1111111
m919.pep	310 KLGQTSMQGIKSYM	:		111:11:11		
m919.pep	370 VDRHYITLGAPLFV !!!!!!!!!!! VDRHYITLGAPLFV 370	111111111	1111111111	1111111111		111111
m919.pep	430 QKTTGYVWQLLPNGI           QKTTGYVWQLLPNGI 430 440					

### 121 and 121-1

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 3075>: m121.seq

1	ATGGAAACAC	AGCTTTACAT	CGGCATCATG	TCGGGAACCA	GCATGGACGG
51	GGCGGATGCC	GTACTGATAC	GGATGGACGG	CGGCAAATGG	CTGGGCGCGG
101	AAGGGCACGC	CTTTACCCCC	TACCCCGGCA	GGTTACGCCG	CCAATTGCTG
151	GATTTGCAGG	ACACAGGCGC	AGACGAACTG	CACCGCAGCA	GGATTTTGTC
201	GCAAGAACTC	AGCCGCCTAT	ATGCGCAAAC	CGCCGCCGAA	CTGCTGTGCA
251	GTCAAAACCT	CGCACCGTCC	GACATTACCG	CCCTCGGCTG	CCACGGGCAA
301	ACCGTCCGAC	ACGCGCCGGA	ACACGGTTAC	AGCATACAGC	TTGCCGATTT
351	GCCGCTGCTG	GCGxxxxxxx		xxxxxxxxx	
401	XXXXXXXXX	XXXXXXXXX		xxxxxxxxx	
451	XXXXXXXXX	XXXXXXXXX			XXXXXXXXX
501	XXXXXXXXX	xxxxxxxxx	xxxxxxxxx		XXXXXXXXX
551	XXXXXXXXX	XXXXXXXXX	XXXXXXXXX		XXXXXXXXX
601	xxxxxxCAGC	TTCCTTACGA	CAAAAACGGT	GCAAAGTCGG	
651	CATATTGCCG	CAACTGCTCG	ACAGGCTGCT	CGCCCACCCG	TATTTCGCAC
701	AACGCCACCC	TAAAAGCACG	GGGCGCGAAC		AAATTGGCTC
751	GAAACCTACC	TTGACGGCGG	CGAAAACCGA		TGCGGACGCT
801	TTCCCGTTTT	ACCGCGCAAA			
851	CAGATGCCCG	TCAAATGTAC	ATTTGCGACG	GCGGCATCCG	CAATCCTGTT
901	TTAATGGCGG	ATTTGGCAGA	ATGTTTCGGC	ACACGCGTTT	CCCTGCACAG
951	CACCGCCGAC	CTGAACCTCG	ATCCGCAATG	GGTGGAAGCC	GCCGnATTTG
1001	CGTGGTTGGC	GGCGTGTTGG	ATTAATCGCA	TTCCCGGTAG	TCCGCACAAA

1051 GCAACCGGCG CATCCAAACC GTGTATTCTG AnCGCGGGAT ATTATTATTG

## This corresponds to the amino acid sequence <SEQ ID 3076; ORF 121>: m121.pep

# The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 3077>: g121.seq

```
ATGGAAACAC AGCTTTACAT CGGCATTATG TCGGGAACCA GTATGGACGG
      GGCGGATGCC GTGCTGGTAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
  51
 101 AAGGGCACGC CTTTACCCCC TACCCTGACC GGTTGCGCCG CAAATTGCTG
 151 GATTTGCAGG ACACAGGCAC AGACGAACTG CACCGCAGCA GGATGTTGTC
      GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
     GTCAAAACCT CGCTCCGTGC GACATTACCG CCCTCGGCTG CCACGGGCAA
 251
 301 ACCGTCCGAC ACGCGCCGGA ACACGGTtac AGCATACAGC TTGCCGATTT
     GCCGCTGCTG GCGGAACTGa cgcggatttT TACCGTCggc gacttcCGCA
 401 GCCGCGACCT TGCTGCCGGC GGacaAGGTG CGCCGCTCGT CCCCGCCTTT
 451 CACGAAGCCC TGTTCCGCGA TGACAGGGAA ACACGCGTGG TACTGAACAT
      CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCGGCGCA CCCGCCTTCG
 551 GCTTCGACAC AGGGCCGGGC AATATGCTGA TGGAcgcgtg gacgcaggca
 601 cacTGGcagc TGCCTTACGA CAAAAacggt gcAAAGgcgg cacAAGGCAA
651 catatTGCcg cAACTGCTCG gcaggctGCT CGCCcaccCG TATTTCTCAC
 701 AACCCcaccc aaAAAGCACG GGgcGCGaac TgtttgcccT AAattggctc
     gaaacctAcc ttgacggcgg cgaaaaccga tacgacgtat tgcggacgct
      ttcccgattc accgcgcaaA ccgTttggga cgccgtctca CACGCAGCGG
 801
851 CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
 901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
     CACCGCCGAA CTGAACCTCG ATCCTCAATG GGTGGAGGCG gccgCATTtg
     cgtggttggC GGCGTGTTGG ATTAACCGCA TTCCCGGTAG TCCGCACAAA
      GCGACCGGCG CATCCAAACC GTGTATTCTG GGCGCGGGAT ATTATTATTG
1051
1101
```

## This corresponds to the amino acid sequence <SEQ ID 3078; ORF 121.ng>: g121.pep

```
1 METQLYIGIM SGTSMDGADA VLVRMDGGKW LGAEGHAFTP YPDRLRRKLL
51 DLQDTGTDEL HRSRMLSQEL SRLYAQTAAE LLCSQNLAPC DITALGCHGQ
101 TVRHAPEHGY SIQLADLPLL AELTRIFTVG DFRSRDLAAG GQGAPLVPAF
151 HEALFRDDRE TRVVLNIGGI ANISVLPPGA PAFGFDTGPG NMLMDAWTQA
201 HWQLPYDKNG AKAAQGNILP QLLGRLLAHP YFSQPHPKST GRELFALNWL
251 ETYLDGGENR YDVLRTLSRF TAQTVWDAVS HAAADARQMY ICGGGIRNPV
301 LMADLAECFG TRVSLHSTAE LNLDPQWVEA AAFAWLAACW INRIPGSPHK
351 ATGASKPCIL GAGYYY*
```

# ORF 121 shows 73.5% identity over a 366 as overlap with a predicted ORF (ORF121.ng) from N. gonorrhoeae: m121/g121

```
30
                                    40
         {\tt METQLYIGIMSGTSMDGADAVLIRMDGGKWLGAEGHAFTPYPGRLRRQLLDLQDTGADEL}
m121.pep
         METQLYIGIMSGTSMDGADAVLVRMDGGKWLGAEGHAFTPYPDRLRRKLLDLQDTGTDEL
g121
               10
                      20
                             30
                                    40
                                           50
                      80
                             90
                                   100
                                          110
         \tt HRSRILSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPEHGYSIQLADLPLL
m121.pep
```

1.7

g121	HRSRMLSQELSRL	YAQTAAELLC	SQNLAPCDITA	ALGCHGOTVR	HAPEHGYSIO	T.ADT.PT.T.
	70	80	90	100	110	120
	130	140	150	160	170	180
m121.pep	AXXXXXXXXXXXX	XXXXXXXXXX	XXXXXXXXXX	xxxxxxxxxx	(XXXXXXXXXXX	VVVVVV
	1 : :				MANAAAAA.	AAAAAA
g121	AELTRIFTVGDFR	SRDLAAGGOG	APIVPAFHEAT	אומידפתת פדע	II NITCCTANIT	CUI DDC3
_	130	140	150	160	170	
	190	200	210	220	•	180
m121.pep	XXXXXXXXXXXXXXX				230	240
	:	: 1		ZOGNITEPOTE	RLLAHPYFA	
g121	PAFGFDTGPGNMLN	ないない ない はなない i	[]		:	
<b>3</b>	190	200	210	40GN1F50FFG		
	250	260	270	220	230	240
m121.pep				280	290	300
mrrr, pep	GRELFAINWLETYI	DGGENKIDVI	KILSKITAQI			GGIRNPV
g121	CDEIENINGTEMAT		111111111		111111	11111
9.21	GRELFALNWLETYI 250	POGGENKADAI	KTLSRFTAQI	CVWDAVSHAAA		GGIRNPV
		260	270	280	290	300
-101	310	320	330	340	350	360
m121.pep	LMADLAECFGTRVS	LHSTADLNL	PQWVEAAXFA	WLAACWINRI	PGSPHKATGA	ASKPCIL
1.01			11111111111	111111111	111111111	
g121	LMADLAECFGTRVS	LHSTAELNL	PQWVEAAAFA	WLAACWINRI	PGSPHKATGA	ASKPCIL
	310	320	330	340	350	360
-101						
m121.pep	XAGYYYX					
- 101	11111					
g121	GAGYYYX					

## The following partial DNA sequence was identified in N. meningitidis <SEQ ID 3079>:

```
a121.seq
         ATGGAAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
     51 GGCGGATGCC GTACTGATAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
    101 AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTTACGCCG CAAATTGCTG
    151 GATTTGCAGG ACACAGGCGC GGACGAACTG CACCGCAGCA GGATGTTGTC
    201
         GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
         GTCAAAACCT CGCGCCGTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
    251
    301 ACCGTCAGAC ACGCGCCGGA ACACAGTTAC AGCGTACAGC TTGCCGATTT
    351 GCCGCTGCTG GCGGAACGGA CTCAGATTTT TACCGTCGGC GACTTCCGCA
        GCCGCGACCT TGCGGCCGGC GGACAAGGCG CGCCGCTCGT CCCCGCCTTT
    401
    451
        CACGAAGCCC TGTTCCGCGA CGACAGGGAA ACACGCGCGG TACTGAACAT
         CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCGACGCA CCCGCCTTCG
         GCTTCGACAC AGGACCGGGC AATATGCTGA TGGACGCGTG GATGCAGGCA
    551
    601 CACTGGCAGC TTCCTTACGA CAAAAACGGT GCAAAGGCGG CACAAGGCAA
        CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCCACCCG TATTTCGCAC
    701 AACCCCACCC TAAAAGCACG GGGCGCGAAC TGTTTGCCCT AAATTGGCTC
    751
         GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TGCGGACGCT
         TTCCCGATTC ACCGCGCAAA CCGTTTTCGA CGCCGTCTCA CACGCAGCGG
    801
         CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
    851
         TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
    951 CACCGCCGAA CTGAACCTCG ATCCGCAATG GGTAGAAGCC GCCGCGTTCG
   1001
         CATGGATGGC GGCGTGTTGG GTCAACCGCA TTCCCGGTAG TCCGCACAAA
   1051
         GCAACCGGCG CATCCAAACC GTGTATTCTG GGCGCGGGAT ATTATTATTG
   1101
```

## This corresponds to the amino acid sequence <SEQ ID 3080; ORF 121.a>:

al21.pep					
1	METQLYIGIM	SGTSMDGADA	VLIRMDGGKW	LGAEGHAFTP	YPGRI.RRKI.I.
51	DLQDTGADEL	HRSRMLSQEL	SRLYAQTAAE	LLCSONLAPS	DITALCCHGO
101	TVRHAPEHSY	SVQLADLPLL	AERTQIFTVG	DFRSRDLAAG	GOGAPINPAF
151	HEALFRODRE	TRAVLNIGGI	ANISVLPPDA	PAFGFDTGPG	ACMMACIMIN AND A MAIN A
201	HWQLPYDKNG	AKAAQGNILP	QLLDRLLAHP	YFAQPHPKST	GRELFALNWL
251	ETYLDGGENR	YDVLRTLSRF	TAQTVFDAVS	HAAADARQMY	ICGGGIRNPV

BNSDOCID: <WO___9957280A2_l_>

301 LMADLAECFG TRVSLHSTAE LNLDPQWVEA AAFAWMAACW VNRIPGSPHK
351 ATGASKPCIL GAGYYY*

m121/a121 ORFs 121 and 121.a showed a 74.0% identity in 366 aa overlap

	10	20	30	40	50	60
m121.pep	METQLYIGIMSGTS	SMDGADAVLI	RMDGGKWLGA	ECHAFTDVDC	DIDDOLIDIO	
	1111111111111	1   1   1   1   1   1   1	1 1 1 1 1 1 1 1 1 1			
<b>a1</b> 21	METQLYIGIMSGTS	MDGADAVIT	RMDGGKWIGN		1       :	
	10	20	30	EGNAPTPYPG.		
		20	30	40	50	60
	70	80				
m121.pep			90	100	110	120
mrzr.beb	HRSRILSQELSRLY	AOTAAELLC	SQNLAPSDITA	ALGCHGQTVRI	HAPEHGYSIC	LADLPLL
- 1 0 1	, , , , , , , , , , , , , , , , , , , ,					
a121	**************************************	AOTAAELLC	SQNLAPSDITA	LGCHGOTVR	IAPEHSYSVO	יונוווו
	70	80	90	100	110	120
				200	110	120
	130	140	150	160	170	
m121.pep	AXXXXXXXXXXXXX	XXXXXXXXXX	77 <b>7</b> 7777777777	T 00	170	180
• •	1 : :		,vvvvvvvvvv	.AAAAXXXXXXX	XXXXXXXXX	XXXXXXX
a121		DDI BROGGO	. =	:		
u	AERTQIFTVGDFRS	KDLAAGGQGA	APLVPAFHEAL	FRDDRETRAV	'LNIGGIANI	SVLPPDA
	130	140	150	160	170	180
					-	200
	190	200	210	220	230	240
m121.pep	XXXXXXXXXXXXXX	XXXXXXXXOI	PYDKNGAKSA	OGNITI POT T	PTIAUDVEN.	240
a121	PAFGFDTGPGNMLM	, , TOMH COMMAC	:	000777		1
	190	200	FIDANGAKAA	ÖGNITAÖTTD	RLLAHPYFA	)PHPKST
	100	200	210	220	230	240
	0.50					
101	250	260	270	280	290	300
m121.pep	GRELFAINWLETYL:	DGGENRYDVL	RTLSRFTAQT	VCDAVSHAAA	DAROMYTCH	
a121	GRELFALNWLETYLI	DGGENRYDVL	RTLSRFTAOT	(/ΓΠΔΙ/Ομλλ	DADOMATOC	. [ ] ] [ ]
	250	260	270	280	DWKÖMI T.C.G.	
			210	200	290	300
	310	320	220			
m121.pep			330	340	350	360
	LMADLAECFGTRVSI	PUSTADFULD	POWVEAAXFA	WLAACWINRI	PGSPHKATGA	SKPCIL
a121						
alzi		THE THEFT	POWVEAAAFAV	MAACWVNRI	PGSPHKATCA	SKPCTI
	310	320	330	340	350	360
					550	360
m121.pep	XAGYYYX					
	11111					
a121	GAGYYYX					

# Further work revealed the DNA sequence identified in N. meningitidis <SEQ ID 3081>:

1	ATGGAAACAC	AGCTTTACAT	CGGCATCATG	TCGGGAACCA	GCATGGACGG
51	GGCGGATGCC	GTACTGATAC	GGATGGACGG		CTGGGCGCGG
101	AAGGGCACGC	CTTTACCCCC	TACCCCGGCA		CCAATTGCTG
151	GATTTGCAGG	ACACAGGCGC	AGACGAACTG		GGATTTTGTC
201		AGCCGCCTAT	ATGCGCAAAC		
251		CGCACCGTCC	GACATTACCG		CTGCTGTGCA
301	ACCGTCCGAC		ACACGGTTAC	0001000010	CCACGGGCAA
351	GCCGCTGCTG	GCGGAACGGA	CGCGGATTTT		TTGCCGATTT
401	GCCGCGACCT	TGCGGCCGGC			GACTTCCGCA
451	CACGAAGCCC	TGTTCCGCGA	GGACAAGGCG	CGCCACTCGT	CCCCGCCTTT
501	CGGCGGGATT		CAACAGGGAA	ACACGCGCGG	
551	GCTTCGACAC	GCCAACATCA	GCGTACTCCC	CCCCGACGCA	CCCGCCTTCG
601		AGGGCCGGGC	AATATGCTGA	TGGACGCGTG	GACGCAGGCA
651	CACTGGCAGC	TTCCTTACGA	CAAAAACGGT	GCAAAGGCGG	CACAAGGCAA
	CATATTGCCG	CAACTGCTCG	ACAGGCTGCT	~~~	
701	AACCCCACCC		GGGCGCGAAC	TGTTTGCCCT	AAATTGGCTC
751	GAAACCTACC		CGAAAACCGA	TACGACGTAT	TGCGGACGCT
801	TTCCCGTTTT	ACCGCGCAAA	CCGTTTGCGA		
					CHCGCAGCGG

901 T 951 C 1001 C 1051 G 1101 A	AGATGCCCG TCAAATG TAATGGCGG ATTTGGC. ACCGCCGAC CTGAACC GTGGTTGGC GGCGTGT CAACCGGCG CATCCAA.	AGA ATGTTT TCG ATCCGC TGG ATTAAT ACC GTGTAT	CGGC ACACG AATG GGTGG CGCA TTCCC TCTG ANCGC	CGTTT CCC CAAGCC GCCC CGGTAG TCCC CGGGAT ATTA	IGCACAG SNATTTG GCACAAA ATTATTG	
m121-1.pep		_	_			
1 M	ETQLYIGIM SGTSMDG	ADA VLIRMD	GGKW LGAEG	HAFTP YPG	RLRRQLL	
51 D 101 T	LQDTGADEL HRSRILS( VRHAPEHGY SIQLADL	QEL SRLYAQ Pi.i. Afbrbi	TAAE LLCSQ ETVG DEPSD	NLAPS DITA	ALGCHGQ	
151 H	EALFRDNRE TRAVLNI	GGI ANISVL	PPDA PAFGE	DTGPG NMLN	ADAWTOA	
201 H 251 E	WQLPYDKNG AKAAQGN	ILP QLLDRL	LAHP YFAQP	HPKST GREI	FALNWL	
301 L	TYLDGGENR YDVLRTL: MADLAECFG TRVSLHS:	TAD LNLDPO	DAVS HAAAD WVEA AXFAW	ARQMY ICGO	GIRNPV	
351 A	TGASKPCIL XAGY <u>YY</u> *		., ,	much Ink	.rgsrm(	
m121-1/g121	ORFs 121-1 and	121.ng sh	owed a 95.	6% identit	y in 366	aa overlap
	10	20	30	40	50	60
m121-1.pep	METQLYIGIMSGTSMI	OGADAVLIRM	DGGKWLGAEG	HAFTPYPGRI	RROLLDLOD	TGADEL
g121					11:11111	11:111
9121	10	20	30	40	rkkrttotőd.	TGTDEL 60
m121-1.pep	70 HRSRILSQELSRLYAÇ	80 ספסגנזאמערר	90 זמיידת פסמנוא	100	110	120
• •			[11]	1111111111	11111111	111111
g121	HRSRMLSQELSRLYA	QTAAELLCSQ	NLAPCDITAL	GCHGQTVRHA	PEHGYSIQL	ADLPLL
	70	80	90	100	110	120
	130	140	150	160	170	180
m121-1.pep	AERTRIFTVGDFRSRI	DLAAGGQGAP:	LVPAFHEALF	RDNRETRAVI	NIGGIANIS	VLPPDA
g121		IIIIIIIIII D <b>LAAGGOGA</b> P	!	:    :   RDDRETRVVI	 .NTGGTANTS	 VI.DDCD
	130	140	150	160	170	180 ·
	190	200	210	200		
m121-1.pep	PAFGFDTGPGNMLMDA		210 YDKNGAKAAO	220 GNILPOLLDR	230	240 PHPKST
				111111111	1111111:1	111111
g121	PAFGFDTGPGNMLMDA 190	WTQAHWQLP' 200	YDKNGAKAAQ 210	GNILPQLLGR 220		
	150	200	210	220	230	240
101 1	250	260	270	280	290	300
m121-1.pep	GRELFALNWLETYLDO	GENRYDVLR'	rlsrftaqtv 	CDAVSHAAAD	ARQMYICGG	JIRNPV
g121	GRELFALNWLETYLDG	GENRYDVLR	FLSRFTAQTV	WDAVSHAAAD	ARQMYICGG	IIIIII GIRNPV
	250	260	270	280	290	300
	310	320	330	340	350	360
m121-1.pep	LMADLAECFGTRVSLF	ISTADLNLDP(	QWVEAAXFAW	LAACWINRIP	GSPHKATGAS	SKPCII
g121	IMADIAECECTRUSI	:				
y===	LMADLAECFGTRVSLF 310	320	330	LAACWINRIP 340	GSPHKATGAS 350	SKPCIL 360
m121-1.pep	XAGYYYX					
	11111					
g121	GAGYYYX					

The following DNA sequence was identified in N. meningitidis <SEQ ID 3083>: a121-1.seq

1 ATGGAAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG

51	GGCGGATGCC	GTACTGATAC	GGATGGACGG	CGGCAAATGG	CTGGGCGCGG		
101	AAGGGCACGC	CTTTACCCCC					
151	GATTTGCAGG	ACACAGGCGC		CACCGCAGCA			
201	GCAAGAACTC	AGCCGCCTGT	ACGCGCAAAC				
251	GTCAAAACCT	CGCGCCGTCC	GACATTACCG	CGCCGCCGAA			
301	ACCGTCAGAC		ACACAGTTAC	CCCTCGGCTG			
351	GCCGCTGCTG		CTCAGATTTT		TTGCCGATTT		
401	GCCGCGACCT	TGCGGCCGGC			GACTTCCGCA		
451	CACGAAGCCC				CCCCGCCTTT		
501	CGGCGGGATT		CGACAGGGAA	ACACGCGCGG			
551	GCTTCGACAC		GCGTACTCCC		CCCGCCTTCG		
601	CACTGGCAGC		AATATGCTGA	TGGACGCGTG	GATGCAGGCA		
651		TTCCTTACGA	*************	GCAAAGGCGG	CACAAGGCAA		
701	CATATTGCCG	CAACTGCTCG	ACAGGCTGCT	CGCCCACCCG	TATTTCGCAC		
	AACCCCACCC	TAAAAGCACG	GGGCGCGAAC	TGTTTGCCCT	AAATTGGCTC		
751	GAAACCTACC	TTGACGGCGG	CGAAAACCGA	TACGACGTAT	TGCGGACGCT		
801	TTCCCGATTC	ACCGCGCAAA	CCGTTTTCGA	CGCCGTCTCA			
851	CAGATGCCCG	TCAAATGTAC	ATTTGCGGCG	GCGGCATCCG	CAATCCTGTT		
901	TTAATGGCGG	ATTTGGCAGA	ATGTTTCGGC	ACACGCGTTT	CCCTGCACAG		
951	CACCGCCGAA	CTGAACCTCG	ATCCGCAATG	GGTAGAAGCC	GCCGCGTTCG		
1001	CATGGATGGC	GGCGTGTTGG	GTCAACCGCA	TTCCCGGTAG	TCCGCACAAA		
1051	GCAACCGGCG	CATCCAAACC	GTGTATTCTG	GGCGCGGGAT	ATTATTATTG		
1101	A			CCCCCGGGAI	ALIALTATTG		
esponds to the amino acid sequence <spo 101.1<="" 2004,="" id="" ord="" td=""></spo>							

### This corresponds to the amino acid sequence <SEQ ID 3084; ORF 121-1.a>: a121-1.pep

1 METQLYIGIM SGTSMDGADA VLIRMDGGKW LGAEGHAFTP YPGRLRRKLL 51 DLQDTGADEL HRSRMLSQEL SRLYAQTAAE LLCSQNLAPS DITALGCHGQ 101 TVRHAPEHSY SVQLADLPLL AERTQIFTVG DFRSRDLAAG GQGAPLVPAF 151 HEALFRDDRE TRAVLNIGGI ANISVLPPDA PAFGFDTGPG NMLMDAWMQA
201 HWQLPYDKNG AKAAQGNILP QLLDRLLAHP YFAQPHPKST GRELFALNWL
251 ETYLDGGENR YDVLRTLSRF TAQTVFDAVS HAAADARQMY ICGGGIRNPV
301 LMADLAECFG TRVSLHSTAE LNLDPQWVEA AAFAWMAACW VNRIPGSPHK
351 ATGASKPCIL GAGYYY*

m121-1/a121-1 ORFs 121-1 and 121-1.a showed a 96.4% identity in 366 aa overlap

m121-1.pep	10 20 METQLYIGIMSGTSMDGADAVI                  METQLYIGIMSGTSMDGADAVI 10 20		11111111	LILIE LILIE	
m121-1.pep	70 80 HRSRILSQELSRLYAQTAAELL     :		3   3   1   1   1   1		
m121-1.pep	130 140 AERTRIFTVGDFRSRDLAAGGQ     :                  AERTQIFTVGDFRSRDLAAGGQ 130 140	1 [ 1   3   4   4   4   4   4	•		180 SVLPPDA
ml21-1.pep	190 200 PAFGFDTGPGNMLMDAWTQAHW		11111111		11111
m121-1.pep	250 260 GRELFALNWLETYLDGGENRYDY		1 1 1 1 1 1 1 1		

- 7

```
310
                        320
                               330
                                       340
          LMADLAECFGTRVSLHSTADLNLDPQWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL
m121-1.pep
          LMADLAECFGTRVSLHSTAELNLDPQWVEAAAFAWMAACWVNRIPGSPHKATGASKPCIL
a121
                               330
                                       340
                                               350
m121-1.pep
          XAGYYYX
           111111
a121
          GAGYYYX
```

#### 128 and 128-1

## The following partial DNA sequence was identified in N. meningitidis <SEQ ID 3085>: m128.seq (partial)

```
m128.seq
         ATGACTGACA ACGCACTGCT CCATTTGGGC GAAGAACCCC GTTTTGATCA
         AATCAAAACC GAAGACATCA AACCCGCCCT GCAAACCGCC ATCGCCGAAG
     101 CGCGCGAACA AATCGCCGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA
     151 AACACTGTCG AACCCCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
     201 GGGCGTGGTG TCGCACCTCA ACTGCGTCGC CGACACGCCC GAACTGCGCG
    251 CCGTCTATAA CGAACTGATG CCCGAAATCA CCGTCTTCTT CACCGAAATC
    301 GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
     351 CGAATTCGAC ACCCTCTCCC CCGCACAAAA AACCAAACTC AACCAC
         TACGCCAGCG AAAAACTGCG CGAAGCCAAA TACGCGTTCA GCGAAACCGA
         WGTCAAAAAA TAYTTCCCYG TCGGCAAWGT ATTAAACGGA CTGTTCGCCC
     51
    101 AAMTCAAAAA ACTMTACGGC ATCGGATTTA CCGAAAAAAC YGTCCCCGTC
         TGGCACAAAG ACGTGCGCTA TTKTGAATTG CAACAAAACG GCGAAMCCAT
    151
    201 AGGCGGCGTT TATATGGATT TGTACGCACG CGAAGGCAAA CGCGGCGGCG
    251 CGTGGATGAA CGACTACAAA GGCCGCCGCC GTTTTTCAGA CGGCACGCTG
    301 CAAYTGCCCA CCGCCTACCT CGTCTGCAAC TTCGCCCCAC CCGTCGGCGG
    351 CAGGGAAGCC CGCYTGAGCC ACGACGAAAT CCTCATCCTC TTCCACGAAA
    401 CCGGACACGG GCTGCACCAC CTGCTTACCC AAGTGGACGA ACTGGGCGTA
    451 TCCGGCATCA ACGGCGTAKA ATGGGACGCG GTCGAACTGC CCAGCCAGTT
    501 TATGGAAAAT TTCGTTTGGG AATACAATGT CTTGGCACAA mTGTCAGCCC
    551 ACGAAGAAAC CGGCGTTCCC YTGCCGAAAG AACTCTTSGA CAAAWTGCTC
    601 GCCGCCAAAA ACTTCCAASG CGGCATGTTC yTsGTCCGGC AAWTGGAGTT
         CGCCCTCTTT GATATGATGA TTTACAGCGA AGACGACGAA GGCCGTCTGA
        AAAACTGGCA ACAGGTTTTA GACAGCGTGC GCAAAAAAGT CGCCGTCATC
         CAGCCGCCCG AATACAACCG CTTCGCCTTG AGCTTCGGCC ACATCTTCGC
        AGGCGGCTAT TCCGCAGCTN ATTACAGCTA CGCGTGGGCG GAAGTATTGA
    801
        GCGCGGACGC ATACGCCGCC TTTGAAGAAA GCGACGATGT CGCCGCCACA
    851
    901 GGCAAACGCT TTTGGCAGGA AATCCTCGCC GTCGGGGAT CGCGCAGCGG
    951 NGCAGAATCC TTCAAAGCCT TCCGCGGCCG CGAACCGAGC ATAGACGCAC
   1001 TCTTGCGCCA CAGCGGTTTC GACAACGCGG TCTGA
```

## This corresponds to the amino acid sequence <SEQ ID 3086; ORF 128>: m128.pep (partial)

```
1 MTDNALLHLG EEPRFDQIKT EDIKPALQTA IAEAREQIAA IKAQTHTGWA
51 NTVEPLTGIT ERVGRIWGVV SHLNCVADTP ELRAVYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFD TLSPAQKTKL NH

1 YASEKLREAK YAFSETXVKK YFPVGXVLNG LFAQXKKLYG IGFTEKTVPV
51 WHKDVRYXEL QQNGEXIGGV YMDLYAREGK RGGAWMNDYK GRRFFSDGTL
101 QLPTAYLVCN FAPPVGGREA RLSHDEILIL FHETGHGLHH LLTQVDELGV
151 SGINGVXWDA VELPSQFMEN FVWEYNVLAQ XSAHEETGVP LPKELXDKXL
201 AAKNFQXGMF XVRQXEFALF DMMIYSEDDE GRLKNWQQVL DSVRKKVAVI
251 QPPEYNRFAL SFGHIFAGGY SAAXYSYAWA EVLSADAYAA FEESDDVAAT
301 GKRFWQEILA VGXSRSGAES FKAFRGREPS IDALLRHSGF DNAV*
```

BNSDOCID: <WO___9957280A2_I_>

## The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 3087>:

atgattgaca acgCActgct ccacttgggc gaagaaccCC GTTTTaatca 1 51 aatccaaacc gaagACAtca AACCCGCCGT CCAAACCGCC ATCGCCGAAG CGCGCGGACA AATCGCCGCC GTCAAAGCGC AAACGCACAC CGGCTGGGCG 101 151 AACACCGTCG AGCGTCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG 201 GGGCGTCGTG TCCCATCTCA ACTCCGTCGT CGACACGCCC GAACTGCGCG 251 CCGTCTATAA CGAACTGATG CCTGAAATCA CCGTCTTCTT CACCGAAATC 301 GGACAAGACA TCGAACTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC 351 CGAATTTGCA ACGCTTTCCC CCGCACAAAA AACCAAGCTC GATCACGACC 401 TGCGCGATTT CGTATTGAGC GGCGCGGAAC TGCCGCCCGA ACGGCAGGCA 451 GAACTGGCAA AACTGCAAAC CGAAGGCGCG CAACTTTCCG CCAAATTCTC 501 CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTTAC TTTGACGATG CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCC GCCGCGCAAA GCGAAGGCAA AACAGGTTAC AAAATCGGCT TGCAGATTCC GCACTACCTT GCCGTTATCC AATACGCCGG CAACCGCGAA CTGCGCGAAC AAATCTACCG CGCCTACGTT ACCCGTGCCA GCGAACTTTC AAACGACGGC 701 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCATTGAA 801 AACCGccaaa cTGCTCGGCT TTAAAAATTA CGCCGAATTG TCGCTGGCAA 851 CCAAAATGGC GGACACGCCC GAACAGGTTT TAAACTTCCT GCACGACCTC 901 GCCCGCCGCG CCAAACCCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC 951 CTTCGCCCGC GAACACCTCG GTCTCGCCGA CCCGCAGCCG TGGGACTTGA 1001 GCTACGCCGG CGAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC 1051 GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTTCTGGCAG GCCTGTTCGC 1101 CCAAATCAAA AAACTCTACG GCATCGGATT CGCCGAAAAA ACCGTTCCCG 1151 TCTGGCACAA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCAAAACC 1201 ATCGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG 1251 CGCGTGGATG AACGACtaca AAGGCCGCG CCGCTTTGCC GACGGCGGCGCCTAC CTCGTCTGCA ACTTCGCGC GACGGCCGCC TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCGCCCC GCCCGTCGGC 1351 GGCAAAGAAG CGCGTTTAAG CCACGACGAA ATCCTCACCC TCTTCCACGA 1401 AacCGGCCAC GGACTGCACC ACCTGCTTAC CCAAGTGGAC GAACTGGGCG TGTCCGGCAT CAACGGCGTA GAATGGGACG CGGTCGAACT GCCCAGCCAG 1451 TTTATGGAAA ACTTCGTTTG GGAATACAAT GTATTGGCAC AAATGTCCGC 1501 1551 CCACGAAGAA ACCGGCGAGC CCCTGCCGAA AGAACTCTTC GACAAAATGC TegecGccaa AAACTTCCAG CGCGGTATGT TCCTCGTCCG GCAAATGGAG TTCGCCCTCT TCGATATGAT GATTTACAGT GAAAGCGACG AATGCCGTCT 1651 GAAAAACTGG CAGCAGGTTT TAGACAGCGT GCGCAAAGAA GTcGCCGTCA 1701 TCCAACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCacatCtTC 1751 1801 GCCGGCGCT ATTCCGCAGG CTATTACAGC TACGCATGGG CCGAAGTCCt CAGCACCGAT GCCTACGCCG CCTTTGAAGA AAGCGACGac gtcGCCGCCA 1851 CAGGCAAACG CTTCTGGCAA GAAAtccttg ccgtcggcgg ctCCCGCAGC 1901 geggeggaat cetteaaage etteegega egegaacega geatagaege 1951 ACTGCTGCGC CAaagcggtT TCGACAACGC gGCttgA 2001

## This corresponds to the amino acid sequence <SEQ ID 3088; ORF 128.ng>: g128.pep

1	MIDNALLHLG	EEPRFNQIQT	EDIKPAVOTA	IAEARGQIAA	VKAOTUTCWA
51	NTVERLTGIT	ERVGRIWGVV	SHLNSVVDTP	ELRAVYNELM	DETTURETET
101	GQDIELYNRF	KTIKNSPEFA	TLSPAOKTKL	DHDLRDFVLS	CARLDDERON
151	ELAKLQTEGA	QLSAKFSONV	LDATDAFGIY	FDDAAPLAGI	PEDALAMFAA
201	AAQSEGKTGY	KIGLQIPHYL	AVIOYAGNRE	LREQIYRAYV	TEDALIAMFAA
251	KFDNTANIDR	TLENALKTAK	LLGFKNYAFI.	SLATKMADTP	
301	ARRAKPYAEK	DLAEVKAFAR	EHI-GI-ADDOD	WDLSYAGEKL	EQVLNFLHDL
351	EVKKYFPVGK	VLAGLFAOIK	KLYGIGEAEK	TVPVWHKDVR	KEAKIAFSET
401	IGGVYMDLYA	REGKRGGAWM	NDYKGPPPFA	DGTLQLPTAY	IFELQUNGKT
451	GKEARLSHDE	ILTLFHETGH	GLHHLLTOWN	ELGVSGINGV	LVCNFAPPVG
501				DKMLAAKNFQ	EWDAVELPSQ
551	FALFDMMIYS	ESDECRLKNW	OOM DEMENT	DIMLAAKNIQ	RGMFLVRQME
601	AGGYSAGYYS		ZZ 1 TO O A KVE	VAVIQPPEYN	KFANSFGHIF
			WINUL PERM	VAATGKRFWQ	EILAVGGSRS

### 651 AAESFKAFRG REPSIDALLR QSGFDNAA*

ORF 128 shows 91.7% identity over a 475 aa overlap with a predicted ORF (ORF 128.ng) from N. gonorrhoeae: m128/g128

-100	10	20	30	40	50	60
g128.pep	MIDNALLHLGEEP	RFNQIQTEDI:	KPAVQTAIAE 	ARGQIAAVK	QTHTGWANT	VERLTGIT
m128	MTDNALLHLGEEP	RFDQIKTEDI:		:   AREOIAAIK	:	/EPI.TGIT
	10	20	30	40	50	60
	70	80	90	100		
g128.pep	ERVGRIWGVVSHL			100 TVFFTEIGOD	110 TELYNREKTI	120 KNSPEFA
	11111111111	1 1:11111			11111111111	
m128	ERVGRIWGVVSHL	NCVADTPELR 80	AVYNELMPEI	TVFFTEIGOD	IELYNRFKT]	KNSPEFD
	70	80	90	100	110	120
	130	140	150	160	170	180
g128.pep	TLSPAQKTKLDHD:	LRDFVLSGAE	LPPERQAELA	KLQTEGAQLS	AKFSQNVLDA	TDAFGIY
m128	:  TLSPAQKTKLNH					
	130					
	//					
g128.pep				340 ekt.deakvae	350 SETEVKKYFP	360
3						
m128			YAS	EKLREAKYAF	SETXVKKYFF	VGXVLNG
				10	20	30
			390	400	410	420
g128.pep	LFAQIKKLYGIGF	AEKTVPVWHKI	OVRYFELOON	GKTIGGVYMD	LYAREGKRGG	AWMNDYK
m128	 LFAQXKKLYGIGF	:                EKTVDVWHKT	MBAAEI OOM	::		
	40	50	FO 60	70	B1AREGRAGG	AWMINDYK 90
g128.pep	430 GRRRFADGTLQLP	140 4 Paylvenfadi	NGGKENDI.GI	160 Joetimi eur	470	480
JF-F					I I I I I I I I I I I I I I I I I I I	
m128	GRRRFSDGTLQLP	TAYLVCNFAPI	PVGGREARLS	DEILILFHE	TGHGLHHLLT	'QVDELGV
	100	110	120	130	140	150
			510 !	520	530	540
g128.pep	SGINGVEWDAVELI	SOFMENTVWE	YNVLAQMSAI	<b>IEETGE</b> PLPK	ELFDKMLAAK	NFQRGMF
m128		SOFMENEVWE	 		 EL VDYVI AAY	NEOVOWE
	160	170	180	190	200	210
	550					
g128.pep	550 LVRQMEFALFDMMI		570 INWOOVI:DSVI	SKENVANTODD SKENVANTODD	590 Evndeancec	600
3 1 1	111 111111		11111111			1111111
m128	XVRQXEFALFDMM]	YSEDDEGRLE	CNWQQVLDSVI	KKVAVIQPP:	EYNRFALSFG	HIFAGGY
	220	230	240	250	260	270
			30 e	540	650	660
g128.pep	SAGYYSYAWAEVLS	TDAYAAFEES	DDVAATGKRI	WOEILAVGG:	SRSAAESFKA	FRGREPS
m128	:           SAAXYSYAWAEVLS	 ADAYAAFEES	DDVAATGKPI	MOETLAVOV	CDCCAPCEER	FRORESC
	280	290	300	310	320	FRGREPS 330

BNSDOCID: <WO___9957280A2_j_>

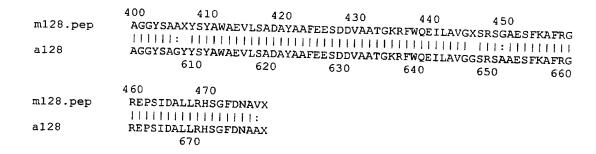
9128.pep | G70 679 | G79 
## The following partial DNA sequence was identified in N. meningitidis <SEQ ID 3089>:

ATGACTGACA ACGCACTGCT CCATTTGGGC GAAGAACCCC GTTTTGATCA AATCAAAACC GAAGACATCA AACCCGCCCT GCAAACCGCC ATTGCCGAAG 51 CGCGCGAACA AATCGCCGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA 151 AACACTGTCG AACCCCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG GGGCGTGGTG TCGCACCTCA ACTCCGTCAC CGACACGCCC GAACTGCGCG CCGCCTACAA TGAATTAATG CCCGAAATTA CCGTCTTCTT CACCGAAATC GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAACTCCCC 301 351 CGAGTTCGAC ACCCTCTCCC ACGCGCAAAA AACCAAACTC AACCACGATC TGCGCGATTT CGTCCTCAGC GGCGCGGAAC TGCCGCCCGA ACAGCAGGCA 401 GAATTGGCAA AACTGCAAAC CGAAGGCGCG CAACTTTCCG CCAAATTCTC 451 CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTTAC TTTGACGATG CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCT 551 601 GCCGCGCAAA GCGAAGGCAA AACAGGCTAC AAAATCGGTT TGCAGATTCC GCACTACCTC GCCGTCATCC AATACGCCGA CAACCGCAAA CTGCGCGAAC AAATCTACCG CGCCTACGTT ACCCGCGCCA GCGAGCTTTC AGACGACGGC 701 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCCCTGCA AACCGCCAAA CTGCTCGGCT TCAAAAACTA CGCCGAATTG TCGCTGGCAA CCAAAATGGC GGACACCCCC GAACAAGTTT TAAACTTCCT GCACGACCTC 851 901 GCCCGCCGCG CCAAACCCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC CTTCGCCCGC GAAAGCCTCG GCCTCGCCGA TTTGCAACCG TGGGACTTGG 951 1001 GCTACGCCGG CGAAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTATTAAACG GACTGTTCGC CCAAATCAAA AAACTCTACG GCATCGGATT TACCGAAAAA ACCGTCCCCG 1101 TCTGGCACAA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCGAAACC **11**51 1201 ATAGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG 1251 CGCGTGGATG AACGACTACA AAGGCCGCCG CCGTTTTTCA GACGGCACGC TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCACCCC GCCCGTCGGC 1351 GGCAAAGAAG CCCGCTTGAG CCATGACGAA ATCCTCACCC TCTTCCACGA 1401 AACCGGACAC GGCCTGCACC ACCTGCTTAC CCAAGTCGAC GAACTGGGCG TATCCGGCAT CAACGGCGTA GAATGGGACG CAGTCGAACT GCCCAGTCAG 1501 TTTATGGAAA ATTTCGTTTG GGAATACAAT GTCTTGGCGC AAATGTCCGC 1551 CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAACTCTTC GACAAAATGC TCGCCGCCAA AAACTTCCAA CGCGGAATGT TCCTCGTCCG CCAAATGGAG 1601 1651 TTCGCCCTCT TTGATATGAT GATTTACAGC GAAGACGACG AAGGCCGTCT GAAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAGAA GTCGCCGTCG 1701 TCCGACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCACATCTTC 1751 GCAGGCGGCT ATTCCGCAGG CTATTACAGC TACGCGTGGG CGGAAGTATT 1801 GAGCGCGGAC GCATACGCCG CCTTTGAAGA AAGCGACGAT GTCGCCGCCA 1851 CAGGCAAACG CTTTTGGCAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC 1901 GCGGCAGAAT CCTTCAAAGC CTTCCGCGGA CGCGAACCGA GCATAGACGC 2001 ACTCTTGCGC CACAGCGGCT TCGACAACGC GGCTTGA

## This corresponds to the amino acid sequence <SEQ ID 3090; ORF 128.a>:

1 MTDNALLHLG EEPRFDQIKT EDIKPALQTA IAEAREQIAA IKAQTHTGWA
51 NTVEPLTGIT ERVGRIWGVV SHLNSVTDTP ELRAAYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFD TLSHAQKTKL NHDLRDFVLS GAELPPEQQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSEGKTGY KIGLQIPHYL AVIQYADNRK LREQIYRAYV TRASELSDDG
251 KFDNTANIDR TLENALQTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
301 ARRAKPYAEK DLAEVKAFAR ESLGLADLQP WDLGYAGEKL REAKYAFSET
351 EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQQNGET
401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTLQLPTAY LVCNFTPPVG
451 GKEARLSHDE ILTLFHETGH GLHHLLTQVD ELGVSGINGV EWDAVELPSQ
501 FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME

551 601 651	FALFDMMIYS EDDEC AGGYSAGYYS YAWAI AAESFKAFRG REPS	EVLSAD AYAA	FEESDD VAA'	VRPPEYN RF/ IGKRFWQ EII	ANSFGHIF LAVGGSRS	
m128/a128 O	RFs 128 and 128.a	showed a 66	5.0% identit	ty in 677 aa	overlan	
m128.pep	10 MTDNALLHLGEE	20 PRFDQIKTEDII	30 KPALOTAIAE	40 AREOIAATKA	50 THTGWANTU	60
a128						111111
	10	20	30	40	50	60
m128.pep	70 ERVGRIWGVVSHI	80 NCVADTPELR	90 AVVNETMPETT	100	110	120
a128	ERVGRIWGVVSHI	1 1:1111	1:11111111		11111111	1111111
2220	70	80	90	100	LLYNRFKTI 110	KNSPEFD 120
m128.pep	130 TLSPAQKTKLNH-					
a128	TLSHAQKTKLNHD 130	LRDFVLSGAEI 140	LPPEQQAELAR 150	(LQTEGAQLSA 160	KFSQNVLDA 170	TDAFGIY 180
m128.pep						
a128	FDDAAPLAGIPED	ALAMFAAAAOS	SEGKTGYKTGI	.ΟΤΡΗΥΙ.ΔΥΤΟ	ישם ואסוועעעע	77VD 7V11
	190	200		220	230	240
m128.pep						
a128	TRASELSDDGKFD	NTANIDRTLEN	IALQTAKLLGF	KNYAELSLAT	KMADTPEQV:	LNFLHDL
	250					
	250	260	270	280	290	300
m128.pep				140	150	
m128.pep a128				140 YASEKLREAK	150 YAFSETXVKI	KYFPVGX
• •				140 YASEKLREAK	150 YAFSETXVKI	KYFPVGX
a128	ARRAKPYAEKDLA 310 160 170	EVKAFARESLG 320	LADLOPWDLG 330	140 YASEKLREAK   :       YAGEKLREAK 340	150 YAFSETXVKI          YAFSETEVKI 350	KYFPVGX        KYFPVGK 360
a128 m128.pep	ARRAKPYAEKDLA 310 160 170 VLNGLFAQXKKLY	EVKAFARESLG 320 180 GIGFTEKTVPV	LADLOPWDLG 330 190 WHKDVRYXEL	140 YASEKLREAK   :       YAGEKLREAK 340 200 QQNGEXIGGV	150 YAFSETXVKI          YAFSETEVKI 350  210 YMDLYAREGE	KYFPVGX        KYFPVGK 360
a128	ARRAKPYAEKDLA 310 160 170 VLNGLFAQXKKLY	EVKAFARESLG 320 180 GIGFTEKTVPV	LADLOPWDLG 330 190 WHKDVRYXEL	140 YASEKLREAK   :       YAGEKLREAK 340 200 QQNGEXIGGV	150 YAFSETXVKI          YAFSETEVKI 350  210 YMDLYAREGE	KYFPVGX        KYFPVGK 360
a128 m128.pep a128	ARRAKPYAEKDLA 310  160 170  VLNGLFAQXKKLY	EVKAFARESLG 320  180 GIGFTEKTVPV           GIGFTEKTVPV 380	ELADLQPWDLG 330 190 WHKDVRYXEL            WHKDVRYFEL 390	140 YASEKLREAK   :       YAGEKLREAK 340 200 QQNGEXIGGV     :     QQNGETIGGV 400	150 YAFSETXVKI           YAFSETEVKI 350  210 YMDLYAREGI          YMDLYAREGI	XYFPVGX        XYFPVGK 360 KRGGAWM        KRGGAWM 420
a128 m128.pep	ARRAKPYAEKDLA 310  160 170 VLNGLFAQXKKLY	EVKAFARESLG 320 180 GIGFTEKTVPV            GIGFTEKTVPV 380 240 LQLPTAYLVCN	ELADLQPWDLG 330  190 WHKDVRYXEL          WHKDVRYFEL 390  250 FAPPVGGREA	140 YASEKLREAK   :       YAGEKLREAK 340 200 QQNGEXIGGV      :     QQNGETIGGV 400 260 RLSHDEILIL	150 YAFSETXVKI           YAFSETEVKI 350 210 YMDLYAREGI          YMDLYAREGI 410 270 FHETGHGLHH	KYFPVGX        KYFPVGK 360  KRGGAWM        KRGGAWM 420
a128 m128.pep a128	ARRAKPYAEKDLA 310  160 170  VLNGLFAQXKKLY	EVKAFARESLG 320 180 GIGFTEKTVPV            GIGFTEKTVPV 380 240 LQLPTAYLVCN	ELADLQPWDLG 330  190 WHKDVRYXEL          WHKDVRYFEL 390  250 FAPPVGGREA	140 YASEKLREAK   :       YAGEKLREAK 340  200 QQNGEXIGGV      :     QQNGETIGGV 400  260 RLSHDEILIL:	150 YAFSETXVKI           YAFSETEVKI 350  210 YMDLYAREGI           YMDLYAREGI 410  270 FHETGHGLHH	KYFPVGX        KYFPVGK 360  KRGGAWM        KRGGAWM 420  HLLTQVD
m128.pep a128 m128.pep	ARRAKPYAEKDLA 310  160 170  VLNGLFAQXKKLY	EVKAFARESLG 320  180 GIGFTEKTVPV            GIGFTEKTVPV 380  240 LQLPTAYLVCN	LADLQPWDLG 330  190 WHKDVRYXEL           WHKDVRYFEL 390  250 FAPPVGGREA  :     :   FTPPVGGKEA	140 YASEKLREAK   :       YAGEKLREAK 340  200 QQNGEXIGGV      :     QQNGETIGGV 400  260 RLSHDEILIL	150 YAFSETXVKI           YAFSETEVKI 350  210 YMDLYAREGI          YMDLYAREGI 410  270 FHETGHGLHI           FHETGHGLHI	KYFPVGX        KYFPVGK 360  KRGGAWM        KRGGAWM 420
m128.pep a128 m128.pep	ARRAKPYAEKDLA 310  160 170  VLNGLFAQXKKLY                             VLNGLFAQIKKLY 370  220 230  NDYKGRRRFSDGT:                              NDYKGRRFSDGT: 430  280 290  ELGVSGINGVXWD	EVKAFARESLG 320  180 GIGFTEKTVPV 380  240 LQLPTAYLVCN           LQLPTAYLVCN 440  300 AVELPSQFMEN	ILADLQPWDLG 330  190 WHKDVRYXEL           WHKDVRYFEL 390  250 FAPPVGGREA  :     :   FTPPVGGKEA 450  310 FVWEYNVLAO	140 YASEKLREAK   :       YAGEKLREAK 340  200 QQNGEXIGGV      :     QQNGETIGGV 400  260 RLSHDEILIL:           460  320 XSAHEETGVP	150 YAFSETXVKI           YAFSETEVKI 350  210 YMDLYAREGI          YMDLYAREGI 410  270 FHETGHGLHI           FHETGHGLHI 470  330	XYFPVGX        XYFPVGK 360  XRGGAWM         XRGGAWM 420  HLLTQVD         480
m128.pep a128 m128.pep a128	ARRAKPYAEKDLA 310  160 170  VLNGLFAQXKKLY	EVKAFARESLG 320  180 GIGFTEKTVPV 380  240 LQLPTAYLVCN 411	ILADLQPWDLG 330  190 WHKDVRYXEL           WHKDVRYFEL 390  250 FAPPVGGREA  :     :   FTPPVGGKEA 450  310 FVWEYNVLAQ	140 YASEKLREAK   :       YAGEKLREAK 340  200 QQNGEXIGGV           QQNGETIGGV 400  260 RLSHDEILIL          460  320 XSAHEETGVP	150 YAFSETXVKI           YAFSETEVKI 350  210 YMDLYAREGI          YMDLYAREGI 410  270 FHETGHGLHI           FHETGHGLHI 470  330 LPKELXDKXI	XYFPVGX        XYFPVGK 360  RRGGAWM         RRGGAWM 420  HLLTQVD         HLLTQVD 480  AAKNFQ
m128.pep a128 m128.pep a128 m128.pep	ARRAKPYAEKDLA 310  160 170 VLNGLFAQXKKLY	EVKAFARESLG 320  180 GIGFTEKTVPV            GIGFTEKTVPV 380  240 LQLPTAYLVCN            LQLPTAYLVCN 440  300 AVELPSQFMEN             AVELPSQFMEN 500	LADLQPWDLG 330  190 WHKDVRYXEL           WHKDVRYFEL 390 250 FAPPVGGREA  :     :   FTPPVGGKEA 450 310 FVWEYNVLAQ            FVWEYNVLAQ	140 YASEKLREAK   :       YAGEKLREAK 340  200 QQNGEXIGGV      :     QQNGETIGGV 400  260 RLSHDEILIL:            RLSHDEILTL: 460  320 XSAHEETGVP:           MSAHEETGVP: 520	150 YAFSETXVKI           YAFSETEVKI 350  210 YMDLYAREGI          YMDLYAREGI 410  270 FHETGHGLHI           FHETGHGLHI 470  330 LPKELXDKXI           LPKELFDKMI 530	KYFPVGX        KYFPVGK 360  KRGGAWM         KRGGAWM 420  HLLTQVD        HLLTQVD 480  AAKNFQ         AAKNFQ 540
m128.pep a128 m128.pep a128 m128.pep	ARRAKPYAEKDLA 310  160 170 VLNGLFAQXKKLY	EVKAFARESLG 320  180 GIGFTEKTVPV            GIGFTEKTVPV 380  240 LQLPTAYLVCN            LQLPTAYLVCN 440  300 AVELPSQFMEN             AVELPSQFMEN 500  360 FDMMIYSEDDE	LADLQPWDLG 330  190 WHKDVRYXEL           WHKDVRYFEL 390 250 FAPPVGGREA  :     :   FTPPVGGKEA 450 310 FVWEYNVLAQ            FVWEYNVLAQ	140 YASEKLREAK   :       YAGEKLREAK 340  200 QQNGEXIGGV      :     QQNGETIGGV 400  260 RLSHDEILIL           RLSHDEILTL  460  320 XSAHEETGVPI 520  380 DSVRKKVAVIO	150 YAFSETXVKI           YAFSETEVKI 350  210 YMDLYAREGI           YMDLYAREGI 410  270 FHETGHGLHI           FHETGHGLHI 470  330 LPKELXDKXI           530  390	KYFPVGX        KYFPVGK 360  KRGGAWM         KRGGAWM 420  HLLTQVD        HLLTQVD 480  AAKNFQ         AAKNFQ 540
m128.pep a128  m128.pep a128  m128.pep a128	ARRAKPYAEKDLA 310  160 170 VLNGLFAQXKKLY	EVKAFARESLG 320  180 GIGFTEKTVPV            GIGFTEKTVPV 380  240 LQLPTAYLVCN            LQLPTAYLVCN 440  300 AVELPSQFMEN              AVELPSQFMEN 500  360 FDMMIYSEDDE	LADLQPWDLG 330  190 WHKDVRYXEL          WHKDVRYFEL 390 250 FAPPVGGREA  :     :   FTPPVGGKEA 450 310 FVWEYNVLAQ            FVWEYNVLAQ	140 YASEKLREAK   :       YAGEKLREAK 340  200 QQNGEXIGGV      :     QQNGETIGGV 400  260 RLSHDEILIL           RLSHDEILTLI 460  320 XSAHEETGVPI 520  380 DSVRKKVAVIC	150 YAFSETXVKI           YAFSETEVKI 350  210 YMDLYAREGI          YMDLYAREGI 410  270 FHETGHGLHI          FHETGHGLHI 470  330 LPKELXDKXI           530  390 QPPEYNRFAI	XYFPVGX         XYFPVGK        XYFPVGK         XRGGAWM           XRGGAWM           420  HLLTQVD         HLLTQVD         AAKNFQ         AAKNFQ



## Further work revealed the DNA sequence identified in N. meningitidis <SEQ ID 3091>:

```
ATGACTGACA ACGCACTGCT CCATTTGGGC GAAGAACCCC GTTTTGATCA
       AATCAAAACC GAAGACATCA AACCCGCCCT GCAAACCGCC ATCGCCGAAG
       CGCGCGAACA AATCGCCGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA
  101
  151 AACACTGTCG AACCCCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
  201 GGGCGTGGTG TCGCACCTCA ACTCCGTCGC CGACACGCCC GAACTGCGCG
  251 CCGTCTATAA CGAACTGATG CCCGAAATCA CCGTCTTCTT CACCGAAATC
  301 GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
       CGAATTCGAC ACCCTCTCCC CCGCACAAAA AACCAAACTC AACCACGATC
  351
       TGCGCGATTT CGTCCTCAGC GGCGCGGAAC TGCCGCCCGA ACAGCAGGCA
  401
      GAACTGGCAA AACTGCAAAC CGAAGGCGCG CAACTTTCCG CCAAATTCTC
  501 CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTTAC TTTGACGATG
      CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCC
  551
      GCCGCGCAAA GCGAAAGCAA AACAGGCTAC AAAATCGGCT TGCAGATTCC
 651 ACACTACCTC GCCGTCATCC AATACGCCGA CAACCGCGAA CTGCGCGAAC
 701 AAATCTACCG CGCCTACGTT ACCCGCGCCA GCGAACTTTC AGACGACGGC
 751 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGCAA ACGCCCTGCA
 801 AACCGCCAAA CTGCTCGGCT TCAAAAACTA CGCCGAATTG TCGCTGGCAA
 851 CCAAAATGGC GGACACGCCC GAACAAGTTT TAAACTTCCT GCACGACCTC
      GCCCGCCGCG CCAAACCCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
 901
 951 CTTCGCCCGC GAAAGCCTGA ACCTCGCCGA TTTGCAACCG TGGGACTTGG
1001 GCTACGCCAG CGAAAAACTG CGCGAAGCCA AATACGCGTT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTATTAAACG GACTGTTCGC
1101 CCAAATCAAA AAACTCTACG GCATCGGATT TACCGAAAAA ACCGTCCCCG
      TCTGGCACAA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCGAAACC
1151
1201 ATAGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
      CGCGTGGATG AACGACTACA AAGGCCGCCG CCGTTTTTCA GACGCACGC
      TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCGCCCC ACCCGTCGGC
1301
      GGCAGGGAAG CCCGCCTGAG CCACGACGAA ATCCTCATCC TCTTCCACGA
1351
      AACCGGACAC GGGCTGCACC ACCTGCTTAC CCAAGTGGAC GAACTGGGCG
1401
      TATCCGGCAT CAACGCCGTA GAATGGGACG CGGTCGAACT GCCCAGCCAG
1451
      TTTATGGAAA ATTTCGTTTG GGAATACAAT GTCTTGGCAC AAATGTCAGC
      CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAACTCTTC GACAAAATGC
      TCGCCGCCAA AAACTTCCAA CGCGGCATGT TCCTCGTCCG GCAAATGGAG
TTCGCCCTCT TTGATATGAT GATTTACAGC GAAGACGACG AAGGCCGTCT
1601
      GAAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAAAA GTCGCCGTCA
1701
1751
      TCCAGCCGCC CGAATACAAC CGCTTCGCCT TGAGCTTCGG CCACATCTTC
      GCAGGCGGCT ATTCCGCAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
1801
     GAGCGCGGAC GCATACGCCG CCTTTGAAGA AAGCGACGAT GTCGCCGCCA
1851
     CAGGCAAACG CTTTTGGCAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
1901
     GCGGCAGAAT CCTTCAAAGC CTTCCGCGGC CGCGAACCGA GCATAGACGC
1951
2001 ACTCTTGCGC CACAGCGGTT TCGACAACGC GGTCTGA
```

# This corresponds to the amino acid sequence <SEQ ID 3092; ORF 128-1>: m128-1.pep.

1 MTDNALLHLG EEPRFDQIKT EDIKPALQTA IAEAREQIAA IKAQTHTGWA 51 NTVEPLTGIT ERVGRIWGVV SHLNSVADTP ELRAVYNELM PEITVFFTEI

```
GQDIELYNRF KTIKNSPEFD TLSPAQKTKL NHDLRDFVLS GAELPPEQQA
               ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
           151
           201 AAQSESKTGY KIGLQIPHYL AVIQYADNRE LREQIYRAYV TRASELSDDG
           251 KFDNTANIDR TLANALQTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
           301 ARRAKPYAEK DLAEVKAFAR ESLNLADLOP WDLGYASEKL REAKYAFSET
          351 EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQQNGET
               IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTLQLPTAY LVCNFAPPVG
               GREARLSHDE ILILFHETGH GLHHLLTQVD ELGVSGINGV EWDAVELPSQ
          451
          501 FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME
          551 FALFDMMIYS EDDEGRLKNW QQVLDSVRKK VAVIQPPEYN RFALSFGHIF
          601 AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRFWQ EILAVGGSRS
          651 AAESFKAFRG REPSIDALLR HSGFDNAV*
The following DNA sequence was identified in N. gonorrhoeae <SEQ ID 3093>:
     g128-1.seq (partial)
            1 ATGATTGACA ACGCACTGCT CCACTTGGGC GAAGAACCCC GTTTTAATCA
              AATCAAAACC GAAGACATCA AACCCGCCGT CCAAACCGCC ATCGCCGAAG
              CGCGCGGACA AATCGCCGCC GTCAAAGCGC AAACGCACAC CGGCTGGGCG
          151 AACACCGTCG AGCGTCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
          201 GGGCGTCGTG TCCCATCTCA ACTCCGTCGT CGACACGCCC GAACTGCGCG
          251
              CCGTCTATAA CGAACTGATG CCTGAAATCA CCGTCTTCTT CACCGAAATC
               GGACAAGACA TCGAACTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
              CGAATTTGCA ACGCTTTCCC CCGCACAAAA AACCAAGCTC GATCACGACC
          351
              TGCGCGATTT CGTATTGAGC GGCGCGGAAC TGCCGCCCGA ACGGCAGGCA
GAACTGGCAA AACTGCAAAC CGAAGGCGCG CAACTTTCCG CCAAATTCTC
          401
              CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTTAC TTTGACGATG
          501
               CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCC
          551
               GCCGCGCAAA GCGAAGGCAA AACAGGTTAC AAAATCGGCT TGCAGATTCC
          601
              GCACTACCTT GCCGTTATCC AATACGCCGG CAACCGCGAA CTGCGCGAAC
          651
          701 AAATCTACCG CGCCTACGTT ACCCGTGCCA GCGAACTTTC AAACGACGGC
          751 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCATTGAA
          801
              AACCGCCAAA CTGCTCGGCT TTAAAAATTA CGCCGAATTG TCGCTGGCAA
              CCAAAATGGC GGACACGCCC GAACAGGTTT TAAACTTCCT GCACGACCTC
          851
          901
              GCCCGCCGCG CCAAACCCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
          951 CTTCGCCCGC GAACACCTCG GTCTCGCCGA CCCGCAGCCG TGGGACTTGA
         1001 GCTACGCCGG CGAAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
         1051 GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTTCTGGCAG GCCTGTTCGC
         1101
               CCAAATCAAA AAACTCTACG GCATCGGATT CGCCGAAAAA ACCGTTCCCG
         1151
               TCTGGCACAA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCAAAACC
               ATCGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
         1201
         1251 CGCGTGGATG AACGACTACA AAGGCCGCCG CCGCTTTGCC GACGGCACGC
         1301 . TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCGCCCC GCCCGTCGGC
         1351 GGCAAAGAAG CGCGTTTAAG CCACGACGAA ATCCTCACCC TCTTCCACGA
               AACCGGCCAC GGACTGCACC ACCTGCTTAC CCAAGTGGAC GAACTGGGCG
         1401
         1451 TGTCCGGCAT CAACGGCGTA AAA
This corresponds to the amino acid sequence <SEQ ID 3094; ORF 128-1.ng>:
     gl28-1.pep (partial)
            1 MIDNALLHLG EEPRFNQIKT EDIKPAVQTA IAEARGQIAA VKAQTHTGWA
              NTVERLTGIT ERVGRIWGVV SHLNSVVDTP ELRAVYNELM PEITVFFTEI
          101 GQDIELYNRF KTIKNSPEFA TLSPAQKTKL DHDLRDFVLS GAELPPERQA
          151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
              AAQSEGKTGY KIGLQIPHYL AVIQYAGNRE LREQIYRAYV TRASELSNDG
          201
          251 KFDNTANIDR TLENALKTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
          301 ARRAKPYAEK DLAEVKAFAR EHLGLADPQP WDLSYAGEKL REAKYAFSET
          351 EVKKYFPVGK VLAGLFAQIK KLYGIGFAEK TVPVWHKDVR YFELQQNGKT
              IGGVYMDLYA REGKRGGAWM NDYKGRRRFA DGTLQLPTAY LVCNFAPPVG
          401
              GKEARLSHDE ILTLFHETGH GLHHLLTQVD ELGVSGINGV K
    m128-1/g128-1
                    ORFs 128-1 and 128-1.ng showed a 94.5% identity in 491 aa
    overlap
                         10
                                    20
                                            30
                                                       40
                                                                 50
                                                                            60
```

BNSDOCID: <WO___9957280A2_I_>

g128-1.pep	MIDNALLH	LGEEPRF	NOIKTEDI	KPAVOTATAE:	ARGQIAAVKA(	O. M. 1	
	, ,,,,,,	1 1 1 1 1 1 1					
m128-1		TOTEL VE	DOTVIEDTI	KPALQTATAE	AREQIAAIKA(	THTGWANTV	EPLTGIT
		10	20	30	40	50	60
		70	80	90	100	110	100
g128-1.pep	ERVGRIWG	VVSHLNS	VVDTPELRA	AVYNETMORTS	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~	TIT ISSES WARE	120 KNSPEFA
m128-1			1 - 1 1 1 1 1 1				
	2111 0111110	70	80	AVYNELMPEIT 90	VFFTEIGQDI	ELYNRFKTI	KNSPEFD
				30	100	110	120
g128-1.pep		30 31 Dubi bi	140	150	160	170	180
g120-1.pep	ILSPACKTE	KLDHDLRI	DFVLSGAEI	PPERQAELAK	LQTEGAQLSA	KFSQNVLDA	TDAFGIY
m128-1	TLSPAQKT	KLNHDLRI	DFVLSGAEL	PPEOOAELAK	LQTEGAQLSA	VECONUL DA	
	13	30	140	150	160	AFSONVLDA 170	TDAFGIY
	19	20	200				100
g128-1.pep	FDDAAPLAG	SIPEDALA	MEAAAAOS	210 EGKTGYKIGI	220 QIPHYLAVIQ	230	240
100 1							
m128-1	FDDAAPLAG	TIBORDA	TIL WWWWO	PRICINITE	QIPHYLAVIQ	YADNRELRE	2IYRAYV
	19	0	200	210	220	230	240
	25		260	270	280	290	300
g128-1.pep	TRASELSND	GKFDNTA	NIDRTLEN	ALKTAKLLGF:	KNVARICIAM		
m128-1		111111		1 4 * 1 1 1 1 1 1 1			
	25	OLUL DIALE	260	ALQTAKLLGF) 270	KNYAELSLATI 280	KMADTPEQVI	NFLHDL
				2.0	280	290	300
g128-1.pep	31		320	330	340	350	360
9120 1.pcp	IIIIIIII	TIIIIII VDTWFAK	AFAREHLG.	LADPQPWDLS	AGEKLREAK	YAFSETEVKK	YFPVGK
m128-1		ALD TIVE A IV	AFARESLN:	LADLQPWDLG	(ASEKLREAK)		VEDUCE
	31	0	320	330	340	350	360
	37	0	380	390	400		
g128-1.pep	VLAGLFAQI	KKLYGIG	FAEKTVPVI	JHKUMDAEET (	ONCRETCOM	410 MDLVADECE	420
m128-1		, , , , , , ,	1 . 1 1 1 1 1 1				
111120-1	VINGLEAQII 370	WITH TOTAL	FTEKTVP <b>V</b> V 380	VHKDVRYFELÇ 390	QNGETIGGVY	MDLYAREGK	RGGAWM
	• • • • • • • • • • • • • • • • • • • •		300	390	400	410	420
~120 1	430	0	440	450	460	470	480
g128-1.pep	NDYKGRRREY	ADGTLQL)	PTAYLVCNE	APPVGGKEAR	LSHDEILTLE	HETGHGLHH	
m128-1	NDYKGRRRFS	DGTLQL	PTAYLVCNE	:    APPVGGREAD			11111
	430	)	440	450	460	470	LLTQVD 480
	490	)					100
g128-1.pep	ELGVSGINGV						
m128-1	ELGVSGINGV	/EWDAVEI	PSQFMENF	VWEYNVLAQM	SAHEETGVPL	PKELFDKML2	AAKNFO
	430	,	300	510	520	530	540
ollowing DNA	sequence v	vas iden	tified in /	V. meninoin	idis <sfo i<="" td=""><td>D 3005~</td><td></td></sfo>	D 3005~	
4+50 1.3ed							
1 ATC	GACTGACA AC	GCACTGC	T CCATTT	GGGC GAAGA	ACCCC GTTT	TGATCA	
	· VERGERALL LIM	CALAN ATT	A DALLOC	וא מכוכרות ותויציצי	~~~~		

## The fo

_	AT GACT GACA	ACGUACTGCT	CCATTTGGGC	GAAGAACCCC	CTTTTCATCA
51	AATCAAAACC	GAAGACATCA	AACCCGCCCT	GCAAACCCCC	ATTITUTE
101	CGCGCGAACA	AATCGCCGCC	ATCAAAGCCC	7777777777	ATTGCCGAAG
151	AACACTGTCG	AACCCCTCAC	CCCCARGO	AAACGCACAC	CGGCTGGGCA
201	AACACTGTCG	TCCCACCTCA	CGGCATCACC	GAACGCGTCG	GCAGGATTTG
251	GGGCGTGGTG	TCGCACCTCA	ACTCCGTCAC	CGACACGCCC	GAACTGCGCG
	CCGCCTACAA	TGAATTAATG	CCCGAAATTA	CCCTCTTCTT	CACCCARAGE
301	GGACAAGACA	TCGAGCTGTA	CAACCGCTTC	$\Delta \Delta \Delta \Delta \Delta C C \Delta T C \Delta$	7777777
351	CGAGITCGAC	ACCCTCTCCC	ACGCGCAAAA	AACCAAACTC	*****
401	TGCGCGATTT	CGTCCTCAGC	GGCGCGGAAC	TCCCCCCCC	AACCACGATC
451	GAATTGGCAA	AACTGCAAAC	CCAACCCCCC	CONTRACTOR	ACAGCAGGCA
		COMMIC	COMMITTEE	CAACTUTECCC	



```
501 CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTTAC TTTGACGATG
      CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCT
     GCCGCGCAAA GCGAAGGCAA AACAGGCTAC AAAATCGGTT TGCAGATTCC
      GCACTACCTC GCCGTCATCC AATACGCCGA CAACCGCAAA CTGCGCGAAC
 701 AAATCTACCG CGCCTACGTT ACCCGCGCCA GCGAGCTTTC AGACGACGGC
      AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCCCTGCA
      AACCGCCAAA CTGCTCGGCT TCAAAAACTA CGCCGAATTG TCGCTGGCAA
 801
     CCAAAATGGC GGACACCCCC GAACAAGTTT TAAACTTCCT GCACGACCTC
 851
 901
     GCCCGCCGCG CCAAACCCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
      CTTCGCCCGC GAAAGCCTCG GCCTCGCCGA TTTGCAACCG TGGGACTTGG
 951
1001
     GCTACGCCGG CGAAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
1051
      GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTATTAAACG GACTGTTCGC
      CCAAATCAAA AAACTCTACG GCATCGGATT TACCGAAAAA ACCGTCCCCG
      TCTGGCACAA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCGAAACC
1151
     ATAGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
1201
      CGCGTGGATG AACGACTACA AAGGCCGCCG CCGTTTTTCA GACGGCACGC
1301
     TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCACCCC GCCCGTCGGC
1351
     GGCAAAGAAG CCCGCTTGAG CCATGACGAA ATCCTCACCC TCTTCCACGA
     AACCGGACAC GGCCTGCACC ACCTGCTTAC CCAAGTCGAC GAACTGGGCG
1401
     TATCCGGCAT CAACGGCGTA GAATGGGACG CAGTCGAACT GCCCAGTCAG
1451
     TTTATGGAAA ATTTCGTTTG GGAATACAAT GTCTTGGCGC AAATGTCCGC
     CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAACTCTTC GACAAAATGC
1551
1601
     TCGCCGCCAA AAACTTCCAA CGCGGAATGT TCCTCGTCCG CCAAATGGAG
     TTCGCCCTCT TTGATATGAT GATTTACAGC GAAGACGACG AAGGCCGTCT
1651
     GAAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAGAA GTCGCCGTCG
1701
     TCCGACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCACATCTTC
1751
     GCAGGCGGCT ATTCCGCAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
     GAGCGCGGAC GCATACGCCG CCTTTGAAGA AAGCGACGAT GTCGCCGCCA
1901
     CAGGCAAACG CTTTTGGCAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
     GCGGCAGAAT CCTTCAAAGC CTTCCGCGGA CGCGAACCGA GCATAGACGC
2001
     ACTCTTGCGC CACAGCGGCT TCGACAACGC GGCTTGA
```

### This corresponds to the amino acid sequence <SEQ ID 3096; ORF 128-1.a>:

a128-1.pep 1 MTDNALLHLG EEPRFDQIKT EDIKPALQTA IAEAREQIAA IKAQTHTGWA NTVEPLTGIT ERVGRIWGVV SHLNSVTDTP ELRAAYNELM PEITVFFTEI GQDIELYNRF KTIKNSPEFD TLSHAQKTKL NHDLRDFVLS GAELPPEQQA ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA AAQSEGKTGY KIGLQIPHYL AVIQYADNRK LREQIYRAYV TRASELSDDG 251 KFDNTANIDR TLENALQTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL 301 ARRAKPYAEK DLAEVKAFAR ESLGLADLQP WDLGYAGEKL REAKYAFSET EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQQNGET IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTLQLPTAY LVCNFTPPVG 401 GKEARLSHDE ILTLFHETGH GLHHLLTQVD ELGVSGINGV EWDAVELPSQ 451 FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME 501 551 FALFDMMIYS EDDEGRLKNW QQVLDSVRKE VAVVRPPEYN RFANSFGHIF AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRFWQ EILAVGGSRS AAESFKAFRG REPSIDALLR HSGFDNAA*

### m128-1/a128-1 ORFs 128-1 and 128-1.a showed a 97.8% identity in 677 aa overlap

	10	20	30	40	50	60
a128-1.pep	MTDNALLHLGEEPF	RFDQIKTEDI	(PALQTAIAE	AREQIAAIKA	OTHTGWANTVE	CPLTGIT
		1111111111				
m128-1	MTDNALLHLGEEPF	FDQIKTEDI	(PALQTAIAE	REQIAAIKA	THTGWANTVE	PLTGTT
	10	20	30	40	50	60
	70	80	90	100	110	120
a128-1.pep	ERVGRIWGVVSHLN	SVTDTPELRA	AYNELMPEIT	VFFTEIGODI	ELYNREKTIK	NSPEED
	1111111111111111	11:111111	:::::::::::::::::::::::::::::::::::::::			111111
m128-1	ERVGRIWGVVSHLN	SVADTPELRA	VYNELMPETT	VEETETGODI	FT.VNDFKTTK	TILLIII Tilliiii
	70	80	90	100	110	120
		- <b>-</b>	,,,	100	110	120
	130	140	150	160	170	180
					•	200

a128-1.pep	TLSHAQKTKLNHDLRDFVLSGAELPPEQQAELAKL	OTEGAOLSAKESONUI DATDAECTV
m128-1		111111111111111111111111111111
11120-1	TLSPAQKTKLNHDLRDFVLSGAELPPEQQAELAKL 130 140 150	QTEGAQLSAKFSQNVLDATDAFGIY 160 170 180
	190 200 210	220 230 240
a128-1.pep	FDDAAPLAGIPEDALAMFAAAAQSEGKTGYKIGLO	IPHYLAVIOYADNRKLREOTYRAYV
m128-1		
	190 200 210	220 230 240
100 1	250 260 270	280 290 300
a128-1.pep	TRASELSDDGKFDNTANIDRTLENALQTAKLLGFK	
m128-1	TRASELSDDGKFDNTANIDRTLANALQTAKLLGFK	NYAELSLATKMADTPEQVLNFLHDL
	2,0	280 290 300
a128-1.pep	310 320 330 ARRAKPYAEKDLAEVKAFARESLGLADLQPWDLGY	340 350 360
m128-1	-	1:111111111111111111111111111
m128-1	ARRAKPYAEKDLAEVKAFARESLNLADLQPWDLGY 310 320 330	ASEKLREAKYAFSETEVKKYFPVGK 340 350 360
	370 380 390	
a128-1.pep	VLNGLFAQIKKLYGIGFTEKTVPVWHKDVRYFELOO	ONGETIGGVYMDI.YAREGKRGGAWM
m128-1		ONGETIGGVYMDIYAREGKRGGAWM
	370 380 390	400 410 420
	430 440 450	460 470 480
a128-1.pep	NDYKGRRRFSDGTLQLPTAYLVCNFTPPVGGKEAR) 	LSHDEILTLFHETGHGLHHLLTQVD
m128-1	NDYKGRRRFSDGTLQLPTAYLVCNFAPPVGGREARI	LSHDEILILFHETGHGLHHLLTQVD
	430 440 450	460 470 480
a128-1.pep	490 500 510 ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMS	520 530 540
		.
m128-1	ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMS 490 500 510	SAHEETGVPLPKELFDKMLAAKNFQ 520 530 540
	550 560 570	
a128-1.pep	RGMFLVRQMEFALFDMMIYSEDDEGRLKNWQQVLDS	580 590 600 SVRKEVAVVRPPEYNRFANSFGHIF
m128-1		SVRKKVAVIODDEVNDENI SECHIE
	550 560 570	580 590 600
	610 620 630	640 650 660
a128-1.pep	AGGYSAGYYSYAWAEVLSADAYAAFEESDDVAATG 	RFWQEILAVGGSRSAAESFKAFRG
m128-1	AGGYSAGYYSYAWAEVLSADAYAAFEESDDVAATG	(RFWQEILAVGGSRSAAESFKAFRG
	610 620 630	640 650 660
a128-1.pep	670 679 REPSIDALLRHSGFDNAAX	
m128-1	111111111111111111111111111111111111111	
m1/8-1	DDDCTD11 thucome	
11120 1	REPSIDALLRHSGFDNAVX 670	

206

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 3097>: m206.seq

k moli o li Mila

```
1 ATGTTTCCCC CCGACAAAAC CCTTTTCCTC TGTCTCAGCG CACTGCTCCT
 51 CGCCTCATGC GGCACGACCT CCGGCAAACA CCGCCAACCG AAACCCAAAC
101 AGACAGTCCG GCAAATCCAA GCCGTCCGCA TCAGCCACAT CGACCGCACA
151 CAAGGCTCGC AGGAACTCAT GCTCCACAGC CTCGGACTCA TCGGCACGCC
    CTACAAATGG GGCGGCAGCA GCACCGCAAC CGGCTTCGAT TGCAGCGGCA
     TGATTCAATT CGTTTACAAr AACGCCCTCA ACGTCAAGCT GCCGCGCACC
301 GCCCGCGACA TGGCGGCGGC AAGCCGSAAA ATCCCCGACA GCCGCYTCAA
351 GGCCGGCGAC CTCGTATTCT TCAACACCGG CGGCGCACAC CGCTACTCAC
401 ACGTCGGACT CTACATCGGC AACGGCGAAT TCATCCATGC CCCCAGCAGC
451 GGCAAAACCA TCAAAACCGA AAAACTCTCC ACACCGTTTT ACGCCAAAAA
501 CTACCTCGGC GCACATACTT TTTTTACAGA ATGA
```

This corresponds to the amino acid sequence <SEQ ID 3098; ORF 206>: m206.pep..

- MFPPDKTLFL CLSALLLASC GTTSGKHROP KPKQTVRQIQ AVRISHIDRT QGSQELMLHS LGLIGTPYKW GGSSTATGFD CSGMIQFVYK NALNVKLPRT 51 101 ARDMAAASRK IPDSRXKAGD LVFFNTGGAH RYSHVGLYIG NGEFIHAPSS
- 151 GKTIKTEKLS TPFYAKNYLG AHTFFTE*

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 3099>: g206.seg

atgitticce cegacaaaac cettiteete tgieteggeg cactgeteet cgcctcatgc ggcacgacct ccggcaaaca ccgccaaccg aaacccaaac 101 agacagteeg geaaateeaa geegteegea teageeacat eggeegeaca 151 caaggetege aggaacteat getecacage eteggactea teggeacgee 201 ctacaaatgg ggcggcagca gcaccgcaac cggcttcgac tgcagcggca 251 tgattcaatt ggtttacaaa aacgccctca acgtcaagct gccgcgcacc 301 gcccgcgaca tggcggcggc aagccgcaaa atccccgaca gccgcctcaa 351 ggccggcgac atcgtattct tcaacaccgg cggcgcacac cgctactcac 401 acgtcggact ctacatcggc aacggcgaat tcatccatgc ccccggcagc 451 ggcaaaacca tcaaaaccga aaaactctcc acaccgtttt acgccaaaaa ctaccttgga gcgcatacgt tttttacaga atga

This corresponds to the amino acid sequence <SEQ ID 3100; ORF 206.ng>: g206.pep

- 1 MFSPDKTLFL CLGALLLASC GTTSGKHRQP KPKQTVRQIQ AVRISHIGRT 51 QGSQELMLHS LGLIGTPYKW GGSSTATGFD CSGMIQLVYK NALNVKLPRT ARDMAAASRK IPDSRLKAGD IVFFNTGGAH RYSHVGLYIG NGEFIHAPGS
- GKTIKTEKLS TPFYAKNYLG AHTFFTE*

ORF 206 shows 96.0% identity over a 177 aa overlap with a predicted ORF (ORF 206.ng) from N. gonorrhoeae: m206/g206

20 30 40 60 MFPPDKTLFLCLSALLLASCGTTSGKHRQPKPKQTVRQIQAVRISHIDRTQGSQELMLHS m206.pep g206 MFSPDKTLFLCLGALLLASCGTTSGKHRQPKPKQTVRQIQAVRISHIGRTQGSQELMLHS 10 20 30 40 50 70 80 90 100 110  ${\tt LGLIGTPYKWGGSSTATGFDCSGMIQFVYKNALNVKLPRTARDMAAASRKIPDSRXKAGD}$ m206.pep g206 LGLIGTPYKWGGSSTATGFDCSGMIQLVYKNALNVKLPRTARDMAAASRKIPDSRLKAGD 70 80 90 100 110 120 130 140 150 160 170 LVFFNTGGAHRYSHVGLYIGNGEFIHAPSSGKTIKTEKLSTPFYAKNYLGAHTFFTEX m206.pep g206 IVFFNTGGAHRYSHVGLYIGNGEFIHAPGSGKTIKTEKLSTPFYAKNYLGAHTFFTE

130 140 150 160 170

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 3101>:

```
ATGTTTCCC CCGACAAAC CCTTTTCCTC TGTCTCAGCG CACTGCTCCT

CGCCTCATGC GGCACACCT CCGGCAAACA CCGCCAACCG AAACCCAAAC

CAAGGCTCGC GCAAATCCAA GCCGTCCGCA TCAGCCACAT CGACCGCACA

CTACAAATGG GGCGCAGCA GCACCGCAAC CGGCTTCGAT TGCAGCGCC

CTACAAATGG GGCGCAGCA GCACCGCAAC CGGCTTCGAT TGCAGCGCCAC

CTACAAATG GGCGCAGCA AACCCCTCA ACGTCAAGCT GCCGCACC

CTGTATTCAAA AACGCCCTCA ACGTCAAGCT GCCGCCACC

ACGCCGCACAC CGCCTCAAA

CTCGTATTCT TCAACACCGG CGCCGCACA CGCCCTTAA

ACGCCGCAAA ATCCCCGACA CGCTTCTCAC

CTATATCGGC AACGCCGAAT TCATCCATGC CCCCCAGCAGC

ACGCCAAAAA

CTACCTCGGC GCACATACTT TCTTTACAGA ATGA
```

This corresponds to the amino acid sequence <SEQ ID 3102; ORF 206.a>:

1 MFPPDKTLFL CLSALLLASC GTTSGKHRQP KPKQTVRQIQ AVRISHIDRT
51 QGSQELMLHS LGLIGTPYKW GGSSTATGFD CSGMIQFVYK NALNVKLPRT
101 ARDMAAASRK IPDSRLKAGD LVFFNTGGAH RYSHVGLYIG NGEFIHAPSS
151 GKTIKTEKLS TPFYAKNYLG AHTFFTE*

m206/a206 ORFs 206 and 206.a showed a 99.4% identity in 177 aa overlap

m206.pep	MEDDDAMI DI GI GA	20	30	40	50	60
mzoo.pep	MFPPDKTLFLCLSA	LLLASCGTTS	SGKHRQPKPK	QTVRQIQAVRI	SHIDRTQGS	QELMLHS
a206	MFPPDKTLFLCLSA	1 [ 1 1 1 1 1 1 [ ]	1	1111111111		
	70	80	90	100	110	120
m206.pep	LGLIGTPYKWGGSS	TATGFDCSGM	IIQFVYKNALI	NVKLPRTARDM	AAASRKIDD	TZU
a206	LGLIGTPYKWGGSS	TATGFDCSGM		1111111111	1111111111	
	70	80	90	100	110	120
m206.pep	130	140	150	160	170	
co.pep	LVFFNTGGAHRYSH	VGTITGNGEF	THAPSSGKT	KTEKLSTPFY.	AKNYLGAHTF	FTEX
a206	LVFFNTGGAHRYSH		  HAPSSGKT]			
	130	140	150	160	170	LIEV

287

## The following partial DNA sequence was identified in N. meningitidis <SEQ ID 3103>:

m∠8/.seq				-	_
1	ATGTTTAAAC	GCAGCGTAAT	CGCAATGGCT	TGTATTTTC	CCCTTTCAGC
51	CTGCGGGGGC	GGCGGTGGCG	GATCGCCCGA	TOTONACTO	GCGGACACGC
101	TGTCAAAACC	TECCECCCCT	GTTCTTTCTC	DOICHAGICG	AGAGGCAAAG
151	CAACATCCCC	CACACCCAC	GIIGIIICIG	AAAAAGAGAC	AGAGGCAAAG
	GAAGAIGCGC	CACAGGCAGG	TTCTCAAGGA	CAGGGCGCGC	CATCCGCACA
201	AGGCAGTCAA	GATATGGCGG	CGGTTTCGGA	AGAAAATACA	CCCAATCCCC
251	GTGCGGTAAC	AGCGGATAAT	CCCAAAAATC	AACACCACCE	COCAMIGGCG
301	GATATGCCCC	AAAATCCCCC	CCCMACAG	AAGACGAGGT	GGCACAAAAT
351	GATATGCCGC	AAAA1GCCGC	CGGTACAGAT	AGTTCGACAC	CGAATCACAC
	CCCGGATCCG	AATATGCTTG	CCGGAAATAT	GGAAAATCAA	CCAACCCAMC
401	CCGGGGGAATC	GTCTCAGCCG	GCAAACCAAC	CCCATATCCC	7 7 THOODERS
451	GACGGAATGC	ACCCCCACCA	TCCCTTCCCTTC	CGGMIMIGGC	AAATGCGGCG
501	GACGGAATGC	AGGGGGACGA	TCCGTCGGCA	GGCGGGCAAA	ATGCCGGCAA
	TACGGCTGCC	CAAGGTGCAA	ATCAAGCCGG	ΔΔΔΓΔΔΠΓΛΛ	CCCCCCCCC
551	CTTCAGATCC	CATCCCCGCG	TCAAACCCTC	CACCMCCCAA	7000000011
			- Character 1 G	CACCIGCGAA	TGGCGGTAGC

```
601 AATTTTGGAA GGGTTGATTT GGCTAATGGC GTTTTGATTG ACGGGCCGTC
 651 GCAAAATATA ACGTTGACCC ACTGTAAAGG CGATTCTTGT AGTGGCAATA
 701 ATTTCTTGGA TGAAGAAGTA CAGCTAAAAT CAGAATTTGA AAAATTAAGT
     GATGCAGACA AAATAAGTAA TTACAAGAAA GATGGGAAGA ATGATAAATT
 801
     TGTCGGTTTG GTTGCCGATA GTGTGCAGAT GAAGGGAATC AATCAATATA
 851
     TTATCTTTA TAAACCTAAA CCCACTTCAT TTGCGCGATT TAGGCGTTCT
 901 GCACGGTCGA GGCGGTCGCT TCCGGCCGAG ATGCCGCTGA TTCCCGTCAA
 951 TCAGGCGGAT ACGCTGATTG TCGATGGGGA AGCGGTCAGC CTGACGGGGC
1001 ATTCCGGCAA TATCTTCGCG CCCGAAGGGA ATTACCGGTA TCTGACTTAC
1051
     GGGGCGGAAA AATTGCCCGG CGGATCGTAT GCCCTTCGTG TTCAAGGCGA
     ACCGGCAAAA GGCGAAATGC TTGCGGGCGC GGCCGTGTAC AACGGCGAAG
1101
     TACTGCATTT CCATACGGAA AACGGCCGTC CGTACCCGAC CAGGGGCAGG
1151
     TTTGCCGCAA AAGTCGATTT CGGCAGCAAA TCTGTGGACG GCATTATCGA
1251
     CAGCGGCGAT GATTTGCATA TGGGTACGCA AAAATTCAAA GCCGCCATCG
     ATGGAAACGG CTTTAAGGGG ACTTGGACGG AAAATGGCAG CGGGGATGTT
     TCCGGAAAGT TTTACGGCCC GGCCGGCGAG GAAGTGGCGG GAAAATACAG
     CTATCGCCCG ACAGATGCGG AAAAGGGCGG ATTCGGCGTG TTTGCCGGCA
1401
1451 AAAAAGAGCA GGATTGA
```

#### This corresponds to the amino acid sequence <SEQ ID 3104; ORF 287>:

```
m287.pep

1 MFKRSVIAMA CIFALSACGG GGGGSPDVKS ADTLSKPAAP VVSEKETEAK
51 EDAPQAGSQG QGAPSAQGSQ DMAAVSEENT GNGGAVTADN PKNEDEVAQN
101 DMPQNAAGTD SSTPNHTPDP NMLAGNMENQ ATDAGESSQP ANQPDMANAA
151 DGMQGDDPSA GGQNAGNTAA QGANQAGNNQ AAGSSDPIPA SNPAPANGGS
201 NFGRVDLANG VLIDGPSQNI TLTHCKGDSC SGNNFLDEEV QLKSEFEKLS
251 DADKISNYKK DGKNDKFVGL VADSVQMKGI NQYIIFYKPK PTSFARFRRS
301 ARSRRSLPAE MPLIPVNQAD TLIVDGEAVS LTGHSGNIFA PEGNYRYLTY
351 GAEKLPGGSY ALRVQGEPAK GEMLAGAAVY NGEVLHFHTE NGRPYPTRGR
401 FAAKVDFGSK SVDGIIDSGD DLHMGTQKFK AAIDGNGFKG TWTENGSGDV
451 SGKFYGPAGE EVAGKYSYRP TDAEKGGFGV FAGKKEQD*
```

## The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 3105>: g287.seq

```
atgtttaaac gcagtgtgat tgcaatggct tgtatttttc ccctttcagc
     ctgtggggc ggcggtggcg gatcgcccga tgtcaagtcg gcggacacgc
 101 cgtcaaaacc ggccgccccc gttgttgctg aaaatgccgg ggaaggggtg
 151 ctgccgaaag aaaagaaaga tgaggaggca gcgggggtg cgccgcaagc
 201 cgatacgcag gacgcaaccg ccggagaagg cagccaagat atggcggcag
 251 tttcggcaga aaatacaggc aatggcggtg cggcaacaac ggacaacccc
 301
     aaaaatgaag acgcggggc gcaaaatgat atgccgcaaa atgccgccga
 351 atccgcaaat caaacaggga acaaccaacc cgccggttct tcagattccg
 401 cccccgcgtc aaaccctgcc cctgcgaatg gcggtagcga ttttggaagg
 451 acgaacgtgg gcaattctgt tgtgattgac ggaccgtcgc aaaatataac
     gttgacccac tgtaaaggcg attcttgtaa tggtgataat ttattggatg
     aagaagcacc gtcaaaatca gaatttgaaa aattaagtga tgaagaaaaa
 601 attaagcgat ataaaaaaga cgagcaacgg gagaattttg tcggtttggt
 651 tgctgacagg gtaaaaaagg atggaactaa caaatatatc atcttctata
 701 cggacaaacc acctactcgt tctgcacggt cgaggaggtc gcttccggcc
 751 gagattccgc tgattcccgt caatcaggcc gatacgctga ttgtggatgg
     ggaagcggtc agcctgacgg ggcattccgg caatatcttc gcgcccgaag
     ggaattaccg gtatctgact tacggggcgg aaaaattgcc cggcggatcg
     tatgccctcc gtgtgcaagg cgaaccggca aaaggcgaaa tgcttgttgg
 901
 951 cacggccgtg tacaacggcg aagtgctgca tttccatatg gaaaacggcc
1001 gtccgtaccc gtccggaggc aggtttgccg caaaagtcga tttcggcagc
1051 aaatctgtgg acggcattat cgacagcggc gatgatttgc atatgggtac
1101 gcaaaaattc aaagccgcca tcgatggaaa cggctttaag gggacttgga
1151
     cggaaaatgg cggcggggat gtttccggaa ggttttacgg cccggccggc
     gaggaagtgg cgggaaaata cagctatcgc ccgacagatg ctgaaaaggg
     cggattcggc gtgtttgccg gcaaaaaaga tcgggattga
```

## This corresponds to the amino acid sequence <SEQ ID 3106; ORF 287.ng>: g287.pep

MFKRSVIAMA CIFPLSACGG GGGGSPDVKS ADTPSKPAAP VVAENAGEGV

51 101 151 201 251	TNVGNSVVID IKRYKKDEQR	MPQNAAESAN GPSQNITLTH ENFVGLVADR	DATAGEGSQD QTGNNQPAGS CKGDSCNGDN VKKDGTNKYI SLTGHSGNIF	SDSAPASNPA LLDEEAPSKS	PANGGSDFGR EFEKLSDEEK
	INVGNSVVID	GPSQNITLTH	CKGDSCNGDN	LLDEEAPSKS	EFEKLSDEEK
	INKINNDEON	ENFVGLVADR	VKKDGTNKYI	IFYTDKPPTR	SARSRRSLPA
301	PIPLIPVOCEDY	DTLIVDGEAV	SLTGHSGNIF	APEGNYRYLT	YGAEKLPGGS
351	KSADCIIDGC	NGEMLVGTAV	YNGEVLHFHM	ENGRPYPSGG	RFAAKVDFGS
	KSVDGIIDSG EEVAGKYSYR	DEDVERGOES	KAALDGNGFK	GTWTENGGGD	VSGRFYGPAG
	VICKISIK	FIDMENGGEG	VEAGKKDRD*		

# m287/g287 ORFs 287 and 287.ng showed a 70.1% identity in 499 aa overlap

m287.pep	10 MFKRSVIAMAC	2.0 I FALSACGGGGG	30 GSPDVKSADTLS	40 SKPAAPVVSE-		49 KETEA
g287	1111111111		1	1111111		
m287.pep	50 60 KEDAPQAGSQG	70 DGAPSAQGSQDM	80 AAVSEENTGNGG	90	100	
g287	1 / 1 / • /	:::      ATAGEGSQDMA 80		1 • 1 • 1 1 1 1 1 1	1 11111	
m287.pep	110 120 DSSTPNHTPDPN	130 IMLAGNMENQATI	140 AGESSQPANQP	150 DMANAADGMO	1.50	169
g287						
m287.pep	170 180 AQGANQAGNNQA	190 AGSSDPIPASNP	200 APANGGSNFGR	210 VDLANGVLIDO	220 GPSQNITL1	229 THCKGDS
g287	nom/or grator	HIIII IIII AGSSDSAPASNP 130 14	APANGGSDFGR'	:::: : :   PNVGNSVVIDO 160	GPSQNITLT 17	THCKGDS
m287.pep	230 240 CSGNNFLDEEVQ	250 LKSEFEKLSDAD	260 KISNYKKDGKNI	270 OKFVGLVADSV	280 /QMKGINQY	289 ZIIFYKP
g287	: : :   : CNGDNLLDEEAP 180	SKSEFEKLSDEE 190 20	VIVKIKKDEÖKI	ENFVGLVADRV 220	:    :  KKDGTNKY 23	IIFYTD
m287.pep	290 300 KPTSFARFRRSA	310 RSRRSLPAEMPL	320 IPVNQADTLIVI	330 OGEAVSLTGHS	340 GNIFAPEG	349 NYRYLT
<b>g2</b> 87	KPPTRSA		IPVNQADTLIVI 260		  GNIFAPEG   280	 NYRYLT 290
m287.pep	350 360 YGAEKLPGGSYA	370 LRVQGEPAKGEMI	380 LAGAAVYNGEVI	390 HFHTENGRPY	400 PTRGRFAA	409 KVDFGS
g287	 YGAEKLPGGSYA: 300	LRVQGEPAKGEMI 310	LVGTAVYNGEVI 320	HFHMENGRPY	:        PSGGRFAA   340	IIIIII KVDFGS 350
m287.pep	410 420 KSVDGIIDSGDDD			1 •     1   1   1   1   1		
g287	KSVDGIIDSGDDI 360	HMGTOKFKAAII 370	GNGFKGTWTEN 380	GGGDVSGRFY	GPAGEEVA	GKYSYR 410
m287.pep	470 480 PTDAEKGGFGVF	489 AGKKEQDX	-			

q287

1111111111111111111111111 PTDAEKGGFGVFAGKKDRDX 420

```
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 3107>:
     a287.seq
```

```
ATGTTTAAAC GCAGTGTGAT TGCAATGGCT TGTATTGTTG CCCTTTCAGC
      CTGTGGGGGC GGCGGTGGCG GATCGCCCGA TGTTAAGTCG GCGGACACGC
  51
      TGTCAAAACC TGCCGCCCCT GTTGTTACTG AAGATGTCGG GGAAGAGGTG
      CTGCCGAAAG AAAAGAAAGA TGAGGAGGCG GTGAGTGGTG CGCCGCAAGC
      CGATACGCAG GACGCAACCG CCGGAAAAGG CGGTCAAGAT ATGGCGGCAG
 201
      TTTCGGCAGA AAATACAGGC AATGGCGGTG CGGCAACAAC GGATAATCCC
 251
     GAAAATAAAG ACGAGGGACC GCAAAATGAT ATGCCGCAAA ATGCCGCCGA
     TACAGATAGT TCGACACCGA ATCACACCCC TGCACCGAAT ATGCCAACCA
     GAGATATGGG AAACCAAGCA CCGGATGCCG GGGAATCGGC ACAACCGGCA
 401
 451
     AACCAACCGG ATATGGCAAA TGCGGCGGAC GGAATGCAGG GGGACGATCC
      GTCGGCAGGG GAAAATGCCG GCAATACGGC AGATCAAGCT GCAAATCAAG
 501
     CTGAAAACAA TCAAGTCGGC GGCTCTCAAA ATCCTGCCTC TTCAACCAAT
 551
 601 CCTAACGCCA CGAATGGCGG CAGCGATTTT GGAAGGATAA ATGTAGCTAA
 651
     TGGCATCAAG CTTGACAGCG GTTCGGAAAA TGTAACGTTG ACACATTGTA
 701
     AAGACAAAGT ATGCGATAGA GATTTCTTAG ATGAAGAAGC ACCACCAAAA
      TCAGAATTTG AAAAATTAAG TGATGAAGAA AAAATTAATA AATATAAAAA
 751
 801
     AGACGAGCAA CGAGAGAATT TTGTCGGTTT GGTTGCTGAC AGGGTAGAAA
 851 AGAATGGAAC TAACAAATAT GTCATCATTT ATAAAGACAA GTCCGCTTCA
 901 TCTTCATCTG CGCGATTCAG GCGTTCTGCA CGGTCGAGGC GGTCGCTTCC
 951 GGCCGAGATG CCGCTGATTC CCGTCAATCA GGCGGATACG CTGATTGTCG
1001
     ATGGGGAAGC GGTCAGCCTG ACGGGGCATT CCGGCAATAT CTTCGCGCCC
     GAAGGGAATT ACCGGTATCT GACTTACGGG GCGGAAAAAT TGTCCGGCGG
1101
     ATCGTATGCC CTCAGTGTGC AAGGCGAACC GGCAAAAGGC GAAATGCTTG
     CGGGCACGGC CGTGTACAAC GGCGAAGTGC TGCATTTCCA TATGGAAAAC
1151
     GGCCGTCCGT CCCCGTCCGG AGGCAGGTTT GCCGCAAAAG TCGATTTCGG
     CAGCAAATCT GTGGACGGCA TTATCGACAG CGGCGATGAT TTGCATATGG
1251
1301
     GTACGCAAAA ATTCAAAGCC GTTATCGATG GAAACGGCTT TAAGGGGACT
     TGGACGGAAA ATGGCGGCGG GGATGTTTCC GGAAGGTTTT ACGGCCCGGC
1401
     CGGCGAAGAA GTGGCGGGAA AATACAGCTA TCGCCCGACA GATGCGGAAA
1451 AGGGCGGATT CGGCGTGTTT GCCGGCAAAA AAGAGCAGGA TTGA
```

# This corresponds to the amino acid sequence <SEQ ID 3108; ORF 287.a>:

```
a287.pep
         MFKRSVIAMA CIVALSACGG GGGGSPDVKS ADTLSKPAAP VVTEDVGEEV
         LPKEKKDEEA VSGAPQADTQ DATAGKGGQD MAAVSAENTG NGGAATTDNP
         ENKDEGPOND MPONAADTDS STPNHTPAPN MPTRDMGNQA PDAGESAOPA
    101
         NQPDMANAAD GMQGDDPSAG ENAGNTADQA ANQAENNQVG GSQNPASSTN
    201
         PNATNGGSDF GRINVANGIK LDSGSENVTL THCKDKVCDR DFLDEEAPPK
         SEFEKLSDEE KINKYKKDEQ RENFVGLVAD RVEKNGTNKY VIIYKDKSAS
    251
         SSSARFRRSA RSRRSLPAEM PLIPVNQADT LIVDGEAVSL TGHSGNIFAP
    351 EGNYRYLTYG AEKLSGGSYA LSVQGEPAKG EMLAGTAVYN GEVLHFHMEN
    401
         GRPSPSGGRF AAKVDFGSKS VDGIIDSGDD LHMGTQKFKA VIDGNGFKGT
         WTENGGGDVS GRFYGPAGEE VAGKYSYRPT DAEKGGFGVF AGKKEQD*
    451
```

m287/a287 ORFs 287 and 287.a showed a 77.2% identity in 501 aa overlap

m287.pep	MFKRS	10 SVIAMACIFA	20	30	40 SKPAAPVVSE-		49
a287	11111				11111111:1		1 - 11
a207	MFARS	10	20	30	SKPAAPVVTEI 40	OVGEEVLPKE 50	KKDEEA 60
	50	60	7.0	80	90	100	100

80

90

100 KEDAPQAGSQGQGAPSAQGSQDMAAVSEENTGNGGAVTADNPKNEDEVAQNDMPQNAAGT m287.pep a287 VSGAPQADTQ--DATAGKGGQDMAAVSAENTGNGGAATTDNPENKDEGPQNDMPQNAADT 70 80 90 100 110

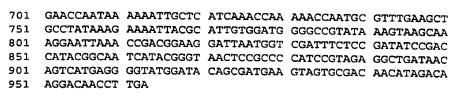
BNSDOCID: <WO 9957280A2 1 >

m287.pep				150 DMANAADGMQ          DMANAADGMQ 160		
m287.pep				210 VDLANGVLID :::   ::  INVANGIKLD 220		
m287.pep	 , , , , , , , ,			270 DKFVGLVADSI :       DFVGLVADRI 280	1	
m287.pep				330 VDGEAVSLTG           VDGEAVSLTG 340		
m287.pep	 . ,,,,,,			390 VLHFHTENGR            VLHFHMENGR 400		
m287.pep a287	 	1 1 1 1 1 1 1 1 1 1	4 1 1 1 1 1 1 1 1	450 ENGSGDVSGK    :    : ENGGGDVSGR 460		
m287.pep a287	480 GGFGVFAGK          GGFGVFAGK 490	HILL				

406

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 3109>:

1	ATGCAAGCAC	GGCTGCTGAT	ACCTATTCTT	TTTTCAGTTT	TTATTTTATC
51	CGCCTGCGGG	ACACTGACAG	GTATTCCATC		GGTAAACGCT
101	TTGCGGTCGA	ACAAGAACTT	GTGGCCGCTT		TGCCGTTAAA
151	GACATGGATT	TACAGGCATT	ACACGGACGA	AAAGTTGCAT	TOTACATEGO
201	CACTATGGGC	GACCAAGGTT	CAGGCAGTTT		CGCTACTCCA
251	TTGATGCACT	GATTCGTGGC	GAATACATAA	ACAGCCCTGC	CCTCCCA
301	GATTACACCT	ATCCACGTTA	CGAAACCACC	GCTGAAACAA	CATCAGAGG
351	TTTGACAGGT	TTAACCACTT	CTTTATCTAC	ACTTAATCAA	CATCAGGCGG
401	CTCGCACCCA	ATCAGACGGT	AGCGGAAGTA	AAAGCACTCT	CCTGCACTCT
451	ATTGGCGGGA	TGGGGGATTA	TCGAAATGAA	ACCOUNTED ACCO	GGGCTTAAAT
501	CGACACTGCC	TTTCTTTCCC	ACTTGGTACA	CACCOCTA	
551	GCATAGACGT				
601	ATCGACGTAT		GCCAATGCCG	ATACAGATGT	GTTTATTAAC
651		CTGAAAGCCC	ACGCAACAGA	ACCGAAATGC	ACCTATACAA
			WARE CAMPACITY	GGAATATTTC	<b>GCACTNCNCN</b>



This corresponds to the amino acid sequence <SEQ ID 3110; ORF 406>: m406.pep

```
1 MOARLLIPIL FSVFILSACG TLTGIPSHGG GKRFAVEQEL VAASARAAVK
51 DMDLQALHGR KVALYIATMG DQGSGSLTGG RYSIDALIRG EYINSPAVRT
101 DYTYPRYETT AETTSGGLTG LTTSLSTLNA PALSRTQSDG SGSKSSLGLN
151 IGGMGDYRNE TLTTNPRDTA FLSHLVQTVF FLRGIDVVSP ANADTDVFIN
201 IDVFGTIRNR TEMHLYNAET LKAQTKLEYF AVDRTNKKLL IKPKTNAFEA
251 AYKENYALWM GPYKVSKGIK PTEGLMVDFS DIRPYGNHTG NSAPSVEADN
301 SHEGYGYSDE VVRQHRQGOP *
```

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 3111>: g406.seq

```
ATGCGGGCAC GGCTGCTGAT ACCTATTCTT TTTTCAGTTT TTATTTTATC
  1
 51 CGCCTGCGGG ACACTGACAG GTATTCCATC GCATGGCGGA GGCAAACGCT
    TCGCGGTCGA ACAAGAACTT GTGGCCGCTT CTGCCAGAGC TGCCGTTAAA
    GACATGGATT TACAGGCATT ACACGGACGA AAAGTTGCAT TGTACATTGC
    AACTATGGGC GACCAAGGTT CAGGCAGTTT GACAGGGGGT CGCTACTCCA
    TTGATGCACT GATTCGCGGC GAATACATAA ACAGCCCTGC CGTCCGCACC
301 GATTACACCT ATCCGCGTTA CGAAACCACC GCTGAAACAA CATCAGGCGG
351 TTTGACGGGT TTAACCACTT CTTTATCTAC ACTTAATGCC CCTGCACTCT
401 CGCGCACCCA ATCAGACGGT AGCGGAAGTA GGAGCAGTCT GGGCTTAAAT
451 ATTGGCGGGA TGGGGGATTA TCGAAATGAA ACCTTGACGA CCAACCCGCG
501 CGACACTGCC TTTCTTTCCC ACTTGGTGCA GACCGTATTT TTCCTGCGCG
551 GCATAGACGT TGTTTCTCCT GCCAATGCCG ATACAGATGT GTTTATTAAC
601 ATCGACGTAT TCGGAACGAT ACGCAACAGA ACCGAAATGC ACCTATACAA
651 TGCCGAAACA CTGAAAGCCC AAACAAAACT GGAATATTTC GCAGTAGACA
701 GAACCAATAA AAAATTGCTC ATCAAACCCA AAACCAATGC GTTTGAAGCT
751 GCCTATAAAG AAAATTACGC ATTGTGGATG GGGCCGTATA AAGTAAGCAA
801 AGGAATCAAA CCGACGGAAG GATTGATGGT CGATTTCTCC GATATCCAAC
851 CATACGGCAA TCATACGGGT AACTCCGCCC CATCCGTAGA GGCTGATAAC
901 AGTCATGAGG GGTATGGATA CAGCGATGAA GCAGTGCGAC AACATAGACA
951 AGGGCAACCT TGA
```

This corresponds to the amino acid sequence <SEQ ID 3112; ORF 406>: g406.pep

```
1 MRARLLIPIL FSVFILSACG TLTGIPSHGG GKRFAVEQEL VAASARAAVK
51 DMDLQALHGR KVALYIATMG DQGSGSLTGG RYSIDALIRG EYINSPAVRT
101 DYTYPRYETT AETTSGGLTG LTTSLSTLNA PALSRTQSDG SGSRSLGLN
151 IGGMGDYRNE TLTTNPRDTA FLSHLVQTVF FLRGIDVVSP ANADTDVFIN
201 IDVFGTIRNR TEMHLYNAET LKAQTKLEYF AVDRTNKKLL IKPKTNAFEA
251 AYKENYALWM GPYKVSKGIK PTEGLMVDFS DIQPYGNHTG NSAPSVEADN
301 SHEGYGYSDE AVRQHRQGQP *
```

ORF 406 shows 98.8% identity over a 320 aa overlap with a predicted ORF (ORF406.a) from N. gonorrhoeae:
q406/m406

BNSDOCID: <WO___9957280A2_J_>

	10	20	30	40	50	60
g406.pep	70 KVALYIATMGDQG; 	80 SGSLTGGRYS] 	90 DALIRGEYI	100 NSPAVRTDYTY	110 PRYETTAET	120 TSGGLTG
m406	KVALYIATMGDQG:	SGSLTGGRYSI 80	DALIRGEYI	NSPAVRTDYT)	PRYETTAET	SGGLTG 120
g406.pep	130 LTTSLSTLNAPALS	140 SRTQSDGSGSR	150 SSLGLNIGG	160 MGDYRNETLTT	170 NPRDTAFLSH	180 ILVQTVF
m406	LTTSLSTLNAPALS	SRTQSDGSGSK 140	SSLGLNIGG 150			 ILVQTVF 180
g406.pep	190 FLRGIDVVSPANAL	200 TDVFINIDVF	210 GTIRNRTEMF	220 !LYNAETLKAQ 	230 TKLEYFAVDR	240 TNKKLL
m406	FLRGIDVVSPANAD	200	210	ILYNAETLKAQ 220	TKLEYFAVDR 230	TNKKLL 240
g406.pep	250   IKPKTNAFEAAYKE 	260 <b>NYALWM</b> GPYK 	270 VSKGIKPTEG	280 LMVDFSDIQP	290 YGNHTGNSAP	300 SVEADN
m406	IKPKTNAFEAAYKE 250	NYALWMGPYK 260	SKGIKPTEG	LMVDFSDIRP	YGNHTGNSAP: 290	SVEADN 300
g406.pep	310 SHEGYGYSDEAVRQ           SHEGYGYSDEVVRQ 310					

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 3113>:

0.009					
1	ATGCAAGCAC	GGCTGCTGAT		TTTTCAGTTT	TTATTTTATC
51	CGCCTGCGGG	ACACTGACAG	GTATTCCATC	GCATGGCGGA	GGTAAACGCT
101	TCGCGGTCGA	ACAAGAACTT	GTGGCCGCTT	CTGCCAGAGC	TGCCGTTAAA
151	GACATGGATT	TACAGGCATT	ACACGGACGA	AAAGTTGCAT	TGTACATTGC
201	AACTATGGGC	GACCAAGGTT	CAGGCAGTTT	GACAGGGGGT	
251	TTGATGCACT	GATTCGTGGC	GAATACATAA	ACAGCCCTGC	CGCTACTCCA
301	GATTACACCT	ATCCACGTTA	CGAAACCACC		CGTCCGTACC
351	TTTGACAGGT	TTAACCACTT	CTTTATCTAC	GCTGAAACAA	CATCAGGCGG
401	CGCGCACCCA	ATCAGACGGT		ACTTAATGCC	CCTGCACTCT
451	ATTGGCGGGA	TGGGGGGATTA	AGCGGAAGTA	AAAGCAGTCT	GGGCTTAAAT
501	CGACACTGCC	TTTCTTTCCC	TCGAAATGAA	ACCTTGACGA	CTAACCCGCG
551	GCATAGACGT		ACTTGGTACA	GACCGTATTT	TTCCTGCGCG
601		TGTTTCTCCT	GCCAATGCCG	ATACGGATGT	GTTTATTAAC
651	ATCGACGTAT	TCGGAACGAT	ACGCAACAGA	ACCGAAATGC	ACCTATACAA
	TGCCGAAACA	CTGAAAGCCC	AAACAAAACT	GGAATATTTC	GCAGTAGACA
701	GAACCAATAA	AAAATTGCTC	ATCAAACCAA	AAACCAATGC	GTTTGAAGCT
751	GCCTATAAAG	AAAATTACGC	ATTGTGGATG	GGACCGTATA	AAGTAAGCAA
801	AGGAATTAAA	CCGACAGAAG	GATTAATGGT	CGATTTCTCC	GATATCCAAC
851	CATACGGCAA	TCATATGGGT	AACTCTGCCC	CATCCGTAGA	GGCTGATAAC
901	AGTCATGAGG	GGTATGGATA	CAGCGATGAA	GCAGTGCGAC	GACATAGACA
951	AGGGCAACCT	TGA		JOI TO LOCGAC	GACATAGACA

# This corresponds to the amino acid sequence <SEQ ID 3114; ORF 406.a>:

1	MOARLLIPIL	FSVFILSACG	TLTGTPSHGG	CKDENVEORT	VAASARAAVK
51	DMDIOALUCE	*****		OWENATORE	VAASARAAVK
JI	DUDTOUTUGE	KVALYIATMG	DOGSGSTTGG	RYSTDAT TRC	EYINSPAVRT
101	DVTVDDVDMM	7 Emmoores	- 20000000	MIDIDALIA	SGSKSSLGLN
101	DITTERIETT	ALITSGGLTG	LTTSTSTIND	DAT SDTOSDC	CCOVOOT OT
151	TCCMCDVDND	M7 MM1133		THISKI ÖSDG	SGSKSSTGTN
101	TOGMODIKNE	TLTINPRDTA	FLSHLVOTVE	FLRCIDIMED	ANADTDVFIN
201	TDUECRIDAD	MENUT INTE		THIGIDANSE	ANADIDAFIN
201	TOVEGITRINK	TEMHLINAET	LKAOTKLEYE	AUDDTMEET	IKPKTNAFEA
				* * * PANAMENT ! !	1 V P V 1 V V V V V V

251 2 301 8	AYKENYALWM GPYKVSKGIK PTEGLMVDFS DIQPYGNHMG NSAPSVEADN SHEGYGYSDE AVRRHRQGQP *
m406/a406	ORFs 406 and 406.a showed a 98.8% identity in 320 aa overlap
m406.pep	10 20 30 40 50 60  MQARLLIPILFSVFILSACGTLTGIPSHGGGKRFAVEQELVAASARAAVKDMDLQALHGR
	10 20 30 40 50 60
m406.pep a406	70 80 90 100 110 120  KVALYIATMGDQGSGSLTGGRYSIDALIRGEYINSPAVRTDYTYPRYETTAETTSGGLTG
	70 80 90 100 110 120 130 140 150 160 170 180
m406.pep a406	LTTSLSTLNAPALSRTQSDGSGSKSSLGLNIGGMGDYRNETLTTNPRDTAFLSHLVQTVF
m406.pep	190 200 210 220 230 240 FLRGIDVVSPANADTDVFINIDVFGTIRNRTEMHLYNAETLKAOTKLEYFAVDRTNKKLL
a406	
m406.pep	250 260 270 280 290 300  IKPKTNAFEAAYKENYALWMGPYKVSKGIKPTEGLMVDFSDIRPYGNHTGNSAPSVEADN
m406.pep a406	310 320 SHEGYGYSDEVVRQHRQGQPX                  SHEGYGYSDEAVRRHRQGQPX 310 320

# **EXAMPLE 2**

# Expression of ORF 919

The primer described in Table 1 for ORF 919 was used to locate and clone ORF 919. The predicted gene 919 was cloned in pET vector and expressed in E. coli. The product of protein expression and purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 919-His fusion protein purification. Mice were immunized with the purified 919-His and sera were used for Western blot (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Symbols: M1, molecular weight marker; PP, purified protein, TP, N. meningitidis total protein extract; OMV, N. meningitidis outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the N. meningitidis immunoreactive band (B). These experiments confirm

 that 919 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 919 are provided in Figure 10. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 919 and the amino acid sequence encoded thereby is provided in Example 1.

# **EXAMPLE 3**

# Expression of ORF 279

The primer described in Table 1 for ORF 279 was used to locate and clone ORF 279. The predicted gene 279 was cloned in pGex vector and expressed in E. coli. The product of protein expression and purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 279-GST purification. Mice were immunized with the purified 279-GST and sera were used for Western blot analysis (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Symbols: M1, molecular weight marker; TP, N. meningitidis total protein extract; OMV, N. meningitidis outer membrane vescicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the N. meningitidis immunoreactive band (B). These experiments confirm that 279 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 279 are provided in Figure 11. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, J. Immunol 143:3007; Roberts et al. 1996, AIDS Res Human Retroviruses 12:593; Quakyi et al. 1992, Scand J Immunol Suppl 11:9). The nucleic acid sequence of ORF 279 and the amino acid sequence encoded thereby is provided in Example 1.

#### **EXAMPLE 4**

# Expression of ORF 576 and 576-1

The primer described in Table 1 for ORF 576 was used to locate and clone ORF 576. The predicted gene 576 was cloned in pGex vector and expressed in E. coli. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 576-GST fusion protein purification. Mice were immunized with the purified 576-GST and sera

S. 1985. S. 1985. P. 3

were used for Western blot (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Symbols: M1, molecular weight marker; TP, N. meningitidis total protein extract; OMV, N. meningitidis outer membrane vescicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the N. meningitidis immunoreactive band (B).. These experiments confirm that ORF 576 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 576 are provided in Figure 12. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, J. Immunol 143:3007; Roberts et al. 1996, AIDS Res Human Retroviruses 12:593; Quakyi et al. 1992, Scand J Immunol Suppl 11:9). The nucleic acid sequence of ORF 576 and the amino acid sequence encoded thereby is provided in Example 1.

# **EXAMPLE 5**

# Expression of ORF 519 and 519-1

The primer described in Table 1 for ORF 519 was used to locate and clone ORF 519. The predicted gene 519 was cloned in pET vector and expressed in E. coli. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 519-His fusion protein purification. Mice were immunized with the purified 519-His and sera were used for Western blot (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Symbols: M1, molecular weight marker; TP, N. meningitidis total protein extract; OMV, N. meningitidis outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the N. meningitidis immunoreactive band (B). These experiments confirm that 519 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 519 are provided in Figure 13. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, J. Immunol 143:3007; Roberts et al. 1996, AIDS Res Human Retroviruses 12:593; Quakyi et al. 1992, Scand J Immunol Suppl 11:9). The nucleic acid sequence of ORF 519 and the amino acid sequence encoded thereby is provided in Example 1.

# **EXAMPLE 6**

# Expression of ORF 121 and 121-1

The primer described in Table 1 for ORF 121 was used to locate and clone ORF 121. The predicted gene 121 was cloned in pET vector and expressed in E. coli. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 121-His fusion protein purification. Mice were immunized with the purified 121-His and sera were used for Western blot analysis (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Results show that 121 is a surface-exposed protein. Symbols: M1, molecular weight marker; TP, N. meningitidis total protein extract; OMV, N. meningitidis outer membrane vescicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the N. meningitidis immunoreactive band (B). These experiments confirm that 121 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 121 are provided in Figure 14. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, J. Immunol 143:3007; Roberts et al. 1996, AIDS Res Human Retroviruses 12:593; Quakyi et al. 1992, Scand J Immunol Suppl 11:9). The nucleic acid sequence of ORF 121 and the amino acid sequence encoded thereby is provided in Example 1.

# EXAMPLE 7

# Expression of ORF 128 and 128-1

The primer described in Table 1 for ORF 128 was used to locate and clone ORF 128. The predicted gene 128 was cloned in pET vector and expressed in E. coli. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 128-His purification. Mice were immunized with the purified 128-His and sera were used for Western blot analysis (panel B), FACS analysis (panel C), bactericidal assay (panel D) and ELISA assay (panel E). Results show that 128 is a surface-exposed protein. Symbols: M1, molecular weight marker; TP, N. meningitidis total protein extract; OMV, N. meningitidis outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the N. meningitidis immunoreactive band (B). These experiments confirm that 128 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 128 are provided in Figure 15. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, J.

Immunol 143:3007; Roberts et al. 1996, AIDS Res Human Retroviruses 12:593; Quakyi et al. 1992, Scand J Immunol Suppl 11:9). The nucleic acid sequence of ORF 128 and the amino acid sequence encoded thereby is provided in Example 1.

#### **EXAMPLE 8**

# Expression of ORF 206

The primer described in Table 1 for ORF 206 was used to locate and clone ORF 206. The predicted gene 206 was cloned in pET vector and expressed in E. coli. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 206-His purification. Mice were immunized with the purified 206-His and sera were used for Western blot analysis (panel B). It is worthnoting that the immunoreactive band in protein extracts from meningococcus is 38 kDa instead of 17 kDa (panel A). To gain information on the nature of this antibody staining we expressed ORF 206 in E. coli without the His-tag and including the predicted leader peptide. Western blot analysis on total protein extracts from E. coli expressing this native form of the 206 protein showed a recative band at a position of 38 kDa, as observed in meningococcus. We conclude that the 38 kDa band in panel B) is specific and that anti-206 antibodies, likely recognize a multimeric protein complex. In panel C is shown the FACS analysis, in panel D the bactericidal assay, and in panel E) the ELISA assay. Results show that 206 is a surface-exposed protein. Symbols: M1, molecular weight marker; TP, N. meningitidis total protein extract; OMV, N. meningitidis outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the N. meningitidis immunoreactive band (B). These experiments confirm that 206 is a surfaceexposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 519 are provided in Figure 16. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, J. Immunol 143:3007; Roberts et al. 1996, AIDS Res Human Retroviruses 12:593; Quakyi et al. 1992, Scand J Immunol Suppl 11:9). The nucleic acid sequence of ORF 206 and the amino acid sequence encoded thereby is provided in Example 1.

# **EXAMPLE 9**

# Expression of ORF 287

The primer described in Table 1 for ORF 287 was used to locate and clone ORF 287. The predicted gene 287 was cloned in pGex vector and expressed in E. coli. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 287-GST fusion protein purification. Mice were immunized with the purified 287-GST and sera were used for FACS analysis (panel B), bactericidal assay (panel C), and ELISA assay (panel D). Results show that 287 is a surface-exposed protein. Symbols: M1, molecular weight marker. Arrow indicates the position of the main recombinant protein product (A). These experiments confirm that 287 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 287 are provided in Figure 17. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, J. Immunol 143:3007; Roberts et al. 1996, AIDS Res Human Retroviruses 12:593; Quakyi et al. 1992, Scand J Immunol Suppl 11:9). The nucleic acid sequence of ORF 287 and the amino acid sequence encoded thereby is provided in Example 1.

# **EXAMPLE 10**

# Expression of ORF 406

The primer described in Table 1 for ORF 406 was used to locate and clone ORF 406. The predicted gene 406 was cloned in pET vector and expressed in E. coli. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 406-His fusion protein purification. Mice were immunized with the purified 406-His and sera were used for Western blot analysis (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Results show that 406 is a surface-exposed protein. Symbols: M1, molecular weight marker; TP, N. meningitidis total protein extract; OMV, N. meningitidis outer membrane vescicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the N. meningitidis immunoreactive band (B). These experiments confirm that 406 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 406 are provided in Figure 18. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, J. Immunol 143:3007; Roberts et al. 1996, AIDS Res Human Retroviruses 12:593; Quakyi et al.

1992, Scand J Immunol Suppl 11:9). The nucleic acid sequence of ORF 406 and the amino acid sequence encoded thereby is provided in Example 1.

# **EXAMPLE 11**

Table 2 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 225 among different strains.

Table 2

225 gene variability	y: List of used Neisseria strains
Identification Strains	Source / reference
number	
Group B	
zo01_225 NG6/88	R. Moxon / Seiler et al., 1996
zo02_225 BZ198	R. Moxon / Seiler et al., 1996
zo03_225 NG3/88	R. Moxon / Seiler et al., 1996
zo04_225 297-0	R. Moxon / Seiler et al., 1996
zo05_225 1000	R. Moxon / Seiler et al., 1996
zo06_225 BZ147	R. Moxon / Seiler et al., 1996
zo07_225 BZ169	R. Moxon / Seiler et al., 1996
zo08_225 528	R. Moxon / Seiler et al., 1996
zo09_225 NGP165	R. Moxon / Seiler et al., 1996
zo10_225 BZ133	R. Moxon / Seiler et al., 1996
zo11_225 NGE31	R. Moxon / Seiler et al., 1996
zo12_225 NGF26	R. Moxon / Seiler et al., 1996
zo13_225 NGE28	R. Moxon / Seiler et al., 1996
zo14_225 NGH38	R. Moxon / Seiler et al., 1996
zo15_225 SWZ107	R. Moxon / Seiler et al., 1996
zo16_225 NGH15	R. Moxon / Seiler et al., 1996
zo17_225 NGH36	R. Moxon / Seiler et al., 1996
zo18_225 BZ232	R. Moxon / Seiler et al., 1996
zo19_225 BZ83	R. Moxon / Seiler et al., 1996
zo20_225 44/76	R. Moxon / Seiler et al., 1996
zo21_225 MC58	R. Moxon
zo96_225 2996	Our collection
Group A	
zo22 225 205900	R. Moxon
zo23 225 F6124	R. Moxon
z2491 Z2491	R. Moxon / Maiden et al., 1998
Group C	
zo24 225 90/18311	R. Moxon
zo25 225 93/4286	R. Moxon
2020_220 /3/4200	IV. IATOVOII

BNSDOCID: <WO___9957280A2_I_>

#### **Others**

zo26_225 A22 (group W) R. Moxon / Maiden et al., 1998

zo27_225 E26 (group X) R. Moxon / Maiden et al., 1998

zo28_225 860800 (group Y) R. Moxon / Maiden et al., 1998

zo29_225 E32 (group Z) R. Moxon / Maiden et al., 1998

# Gonococcus

zo32_225 Ng F62

R. Moxon / Maiden et al., 1998

zo33_225 Ng SN4

R. Moxon

fa1090

FA1090

R. Moxon

# References:

Seiler A. et al., Mol. Microbiol., 1996, 19(4):841-856. Maiden et al., Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.

# The amino acid sequences for each listed strain are as follows:

# >FA1090 <SEQ ID 3115>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPVNRAPARRAG NADELIGSAMGLNEQPVLPVNRAPARRAGNADELIGSAMGLLGIAYRYGGTSVSTGFDCS GFMQHIFKRAMGINLPRTSAEQARMGAPVARSELQPGDMVFFRTLGGSRISHVGLYIGNN RFIHAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDPSRFLN*

#### Z2491 <SEQ ID 3116>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRVPARRAGNA DELIGNAMGLNEQPVLPVNRAPARRAGNADELIGNAMGLLGIAYRYGGTSISTGFDCSGF MQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNNRF IHAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDPSRFLN*

# ZO01_225 <SEQ ID 3117>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

# Z002_225 <SEQ ID 3118>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

#### ZO03_225 <SEQ ID 3119>

MDSFFKPAVWAVLWLMFAVRLALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

in Controlling of the Control

ZO04 225 <SEQ ID 3120>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

#### ZO05 225 <SEQ ID 3121>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGSAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

#### ZO06 225 <SEQ ID 3122>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

#### ZO07 225 <SEQ ID 3123>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQFVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

#### ZO08 225 <SEQ ID 3124>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGSAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

# ZO09_225 <SEQ ID 3125>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

#### ZO10 225 <SEQ ID 3126>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

#### ZO11_225 <SEQ ID 3127>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLNEQPVLPVNRAPARRAGNADELIGNAMGLLGIAYRYGGTSVSTGFDCSGF MQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNNRF IHAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDPSRFLN*

# ZO12_225 <SEQ ID 3128>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

ZO13 225 <SEQ ID 3129>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPVNRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFIQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

#### ZO14 225 <SEQ ID 3130>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

# ZO15_225 <SEQ ID 3131>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLLGIAYRYGGTSVSTGFDCS GFMQHIFKRAMGINLPRTSAEQARMGTPVARSELQFGDMVFFRTLGGSRISHVGLYIGNN RFIHAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDPSRFLN*

# ZO16_225 <SEQ ID 3132>

MDSFFKPAVWAVIWIMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

#### ZO17 225 <SEQ ID 3133>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

# ZO18_225 <SEQ ID 3134>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

# ZO19_225 <SEQ ID 3135>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

# ZO20 225 <SEQ ID 3136>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPINRAPARRAGNADELIGSAMGLNEQPVLPVNRVPARRAGNA DELIGNAMGLNEQPVLPVNRAPARRAGNADELIGNAMGLLGIAYRYGGTSVSTGFDCSGF MQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNNRF IHAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDPSRFLN*

# ZO21_225 <SEQ ID 3137>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

#### ZO22_225 <SEQ ID 3138>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR

.

SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

#### ZO23 225 <SEO ID 3139>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

#### ZO24 225 <SEQ ID 3140>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

#### ZO25 225 <SEQ ID 3141>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

#### ZO26 225 <SEQ ID 3142>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

#### ZO27_225 <SEQ ID 3143>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

# ZO28_225 <SEQ ID 3144>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

#### ZO29 225 <SEQ ID 3145>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

#### ZO32 225 <SEQ ID 3146>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPVNRAPARRAG NADELIGSAMGLNEQPVLPVNRAPARRAGNADELIGSAMGLLGIAYRYGGTSVSTGFDCS GFMQHIFKRAMGINLPRTSAEQARMGAPVARSELQPGDMVFFRTLGGSRISHVGLYIGNN RFIHAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDPSRFLN*

#### ZO33_225 <SEQ ID 3147>

MDSFFKPAVWAVLWLMFAVRSALADELTNLLSSREQILRQFAEDEQPVLPVNRAPARRAG NADELIGSAMGLNEQPVLPVNRAPARRAGNADELIGSAMGLLGIAYRYGGTSVSTGFDCS GFMQHIFKRAMGINLPRTSAEQARMGAPVARSELQPGDMVFFRTLGGSRISHVGLYIGNN RFIHAPRTGKNIEITSLSHKYWSGKYAFARRIKKNDPSRFLN*

ZO96 225 <SEQ ID 3148>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR VKKNDPSRFLN*

Figure 19 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 225, further confirming its utility as an antigen for both vaccines and diagnostics.

# **EXAMPLE 12**

Table 3 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 235 among different strains.

Table 3

235 gen	e variabilit	y: List of used Neisseria strains
	· · · · · · · · · · · · · · · · · · ·	y. List of used incisseria strains
Identifica	ation Strains	Reference
number		
G	roup B	
gnmzq01	NG6/88	Seiler et al., 1996
gnmzq02	BZ198	Seiler et al., 1996
gnmzq03	NG3/88	Seiler et al., 1996
gnmzq04	1000	Seiler et al., 1996
gnmzq05	1000	Seiler et al., 1996
gnmzq07	BZ169	Seiler et al., 1996
gnmzq08	528	Seiler et al., 1996
gnmzq09	NGP165	Seiler et al., 1996
gnmzq10	BZ133	Seiler et al., 1996
gnmzq11	NGE31	Seiler et al., 1996
gnmzq13	NGE28	Seiler et al., 1996
gnmzq14		Seiler et al., 1996
gnmzq15	SWZ107	Seiler et al., 1996
gnmzq16	NGH15	Seiler et al., 1996
gnmzq17	NGH36	Seiler et al., 1996
gnmzq18	BZ232	Seiler et al., 1996
gnmzq19	BZ83	Seiler et al., 1996
gnmzq21	MC58	Virji et al., 1992
Gr	oup A	
gnmzq22	205900	Our collection
<u></u>		Our concentral

gnmzq23 F6124 Our collection z2491 Z2491 Maiden et al., 1998

Group C

gnmzq24 90/18311 Our collection gnmzq25 93/4286 Our collection

Others

gnmzq26 A22 (group W) Maiden et al., 1998 gnmzq27 E26 (group X) Maiden et al., 1998 (group Y) Maiden et al., 1998 gnmzq28 860800 gnmzq29 E32 (group Z) Maiden et al., 1998 gnmzq31 N. lactamica Our collection

Gonococcus

gnmzq32 Ng F62 Maiden et al., 1998 gnmzq33 Ng SN4 Our collection

fa1090 FA1090 Dempsey et al. 1991

# References:

Seiler A. et al., Mol. Microbiol., 1996, 19(4):841-856. Maiden R. et al., Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145. Virji M. et al., Mol. Microbiol., 1992, 6:1271-1279 Dempsey J.F. et al., J. Bacteriol., 1991, 173:5476-5486

The amino acid sequences for each listed strain are as follows:

FA1090 <SEQ ID 3149> MKPLILGLAAVLALSACQVRKAPDLDYTSFKESKPASILVVPPLNESPDVNGTWGMLAST AAPISEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS

YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT DRGYQVSKTAAYNLLSPYSRNGILKGPRFVEEQPK*

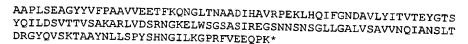
GNMZQ01 <SEQ ID 3150>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST AAPLSEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANNLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEOPK*

GNMZQ02 <SEQ ID 3151>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ03 <SEQ ID 3152> MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST



#### GNMZQ04 <SEQ ID 3153>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

#### GNMZQ05 <SEQ ID 3154>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST AAPLSEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANNLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

#### GNMZQ07 <SEQ ID 3155>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

# GNMZQ08 <SEQ ID 3156>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST AAPLSEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANNLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

#### GNMZQ09 <SEQ ID 3157>

MKPLILGLAAALVLSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGMLAST AEPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVQPEKLHQIFGNDAVLYITITEYGTS YQILDSVTTVSARARLVDSRNGKVLWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

#### GNMZQ10 <SEQ ID 3158>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

#### GNMZQ11 <SEQ ID 3159>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

# GNMZQ13 <SEQ ID 3160>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

#### GNMZQ14 <SEQ ID 3161>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEOPK*

#### GNMZQ15 <SEQ ID 3162>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ16 <SEQ ID 3163>

Committee of the second

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEOPK*

#### GNMZQ17 <SEQ ID 3164>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

# GNMZQ18 <SEQ ID 3165>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

#### GNMZQ19 <SEQ ID 3166>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEOPK*

# GNMZQ21 <SEQ ID 3166>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEOPK*

#### GNM2Q22 <SEQ ID 3167>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEOPK*

#### GNMZQ23 <SEQ ID 3168>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

#### GNMZQ24 <SEQ ID 3169>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

#### GNMZQ25 <SEQ ID 3170>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

#### GNMZQ26 <SEQ ID 3171>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGMLAST AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

# GNMZQ27 <SEQ ID 3172>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEOPK*

BNSDOCID: <WO___9957280A2_I_>

GNMZQ28 <SEQ ID 3173>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ29 <SEQ ID 3174>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEOPK*

GNMZQ31 <SEQ ID 3175>

MKPLILGLAAVLALSACQVQKAPDFDYTAFKESKPASILVVPPLNESPDVNGTWGMLAST AEPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITITEYGTS YQILDSVTTVSARARLVDSRNGKVLWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT DRGYQVSKAAAYDLLSPYSHNGILKGPRFVEEQPK*

GNMZQ32 <SEQ ID 3176>

MKPLILGLAAVLALSACQVRKAPDLDYTSFKESKPASILVVPPLNESPDVNGTWGMLAST AAPISEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT DRGYQVSKTAAYNLLSPYSRNGILKGPRFVEEQPK*

GNMZQ33 <SEQ ID 3177>

MKPLILGLAAVLALSACQVRKAFDLDYTSFKESKPASILVVPPLNESPDVNGTWGMLAST AAPISEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT DRGYQVSKTAAYNLLSPYSRNGILKGPRFVEEQPK*

Z2491 <SEQ ID 3178>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDAVLYITVTEYGTS YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

Figure 20 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 235, further confirming its utility as an antigen for both vaccines and diagnostics.

# **EXAMPLE 13**

Table 4 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 287 among different strains.

Table 4

287 gene variability: List of used Neisseria strains

**Identification Strains** 

Reference

number

. 그 그 시작에도 작는 다음 14년

G	roup B	
287_2	BZ198	Seiler et al., 1996
287_9	NGP165	Seiler et al., 1996
287_14	NGH38	Seiler et al., 1996
287_21	MC58	Virji et al., 1992
G	roup A	
z2491	<b>Z24</b> 91	Maiden et al., 1998
G	onococcus	
fa1090	FA1090	Dempsey et al. 1991

## References:

Seiler A. et al., Mol. Microbiol., 1996, 19(4):841-856. Maiden R. et al., Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145. Virji M. et al., Mol. Microbiol., 1992, 6:1271-1279 Dempsey J.F. et al., J. Bacteriol., 1991, 173:5476-5486

The amino acid sequences for each listed strain are as follows:

#### 287 14 <SEQ ID 3179>

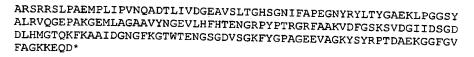
MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSKPAAPVVSEKETEAKEDAPQAGSQG QGAPSAQGGQDMAAVSEENTGNGGAAATDKPKNEDEGAQNDMPQNAADTDSLTPNHTPAS NMPAGNMENQAPDAGESEQPANQPDMANTADGMQGDDPSAGGENAGNTAAQGTNQAENNQ TAGSQNPASSTNPSATNSGGDFGRTNVGNSVVIDGPSQNITLTHCKGDSCSGNNFLDEEV QLKSEFEKLSDADKISNYKKDGKNDGKNDKFVGLVADSVQMKGINQYIIFYKPKPTSFAR FRRSARSRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLTYGAEKLP GGSYALRVQGEPSKGEMLAGTAVYNGEVLHFHTENGRPSPSRGRFAAKVDFGSKSVDGII DSGDGLHMGTQKFKAAIDGNGFKGTWTENGGGDVSGKFYGPAGEEVAGKYSYRPTDAEKG

# 287 2 <SEQ ID 3180>

MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSKPAAPVVSEKETEAKEDAPQAGSQG QGAPSAQGGQDMAAVSEENTGNGGAAATDKPKNEDEGAQNDMPQNAADTDSLTPNHTPAS NMPAGNMENQAPDAGESEQPANQPDMANTADGMQGDDPSAGGENAGNTAAQGTNQAENNQ TAGSQNPASSTNPSATNSGGDFGRTNVGNSVVIDGPSQNITLTHCKGDSCSGNNFLDEEV QLKSEFEKLSDADKISNYKKDGKNDGKNDKFVGLVADSVQMKGINQYIIFYKPKPTSFAR FRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLTYGAEKLP GGSYALRVQGEPSKGEMLAGTAVYNGEVLHFHTENGRPSPSRGRFAAKVDFGSKSVDGII DSGDGLHMGTQKFKAAIDGNGFKGTWTENGGGDVSGKFYGPAGEEVAGKYSYRPTDAEKG

# 287_21. <SEQ ID 3181>

MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSKPAAPVVSEKETEAKEDAPQAGSQG QGAPSAQGSQDMAAVSEENTGNGGAVTADNPKNEDEVAQNDMPQNAAGTDSSTPNHTPDP NMLAGNMENQATDAGESSQPANQPDMANAADGMQGDDPSAGGQNAGNTAAQGANQAGNNQ AAGSSDPIPASNPAPANGGSNFGRVDLANGVLIDGPSQNITLTHCKGDSCSGNNFLDEEV QLKSEFEKLSDADKISNYKKDGKNDKFVGLVADSVQMKGINQYIIFYKPKPTSFARFRRS



287_9 <SEQ ID 3182>

MFKRSVIAMACIVALSACGGGGGGSPDVKSADTLSKPAAPVVTEDVGEEVLPKEKKDEEA VSGAPQADTQDATAGKGGQDMAAVSAENTGNGGAATTDNPENKDEGPQNDMPQNAADTDS STPNHTPAPNMPTRDMGNQAPDAGESAQPANQPDMANAADGMQGDDPSAGENAGNTADQA ANQAENNQVGGSQNPASSTNPNATNGGSDFGRINVANGIKLDSGSENVTLTHCKDKVCDR DFLDEEAPPKSEFEKLSDEEKINKYKKDEQRENFVGLVADRVEKNGTNKYVIIYKDKSAS SSSARFRRSARSRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLTYG AEKLSGGSYALSVQGEPAKGEMLAGTAVYNGEVLHFHMENGRPSPSGGRFAAKVDFGSKS VDGIIDSGDDLHMGTQKFKAVIDGNGFKGTWTENGGGDVSGRFYGPAGEEVAGKYSYRPT DAEKGGFGVFAGKKEQD*

FA1090 <SEQ ID 3183>

MFKRSVIAMACIFPLSACGGGGGGSPDVKSADTPSKPAAPVVAENAGEGVLPKEKKDEEA AGGAPQADTQDATAGEGSQDMAAVSAENTGNGGAATTDNPKNEDAGAQNDMPQNAAESAN QTGNNQPAGSSDSAPASNPAPANGGSDFGRTNVGNSVVIDGPSQNITLTHCKGDSCNGDN LLDEEAPSKSEFEKLSDEEKIKRYKKDEQRENFVGLVADRVKKDGTNKYIIFYTDKPPTR SARSRRSLPAEIPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLTYGAEKLPGGS YALRVQGEPAKGEMLVGTAVYNGEVLHFHMENGRPYPSGGRFAAKVDFGSKSVDGIIDSG DDLHMGTQKFKAAIDGNGFKGTWTENGGGDVSGRFYGPAGEEVAGKYSYRPTDAEKGGFG VFAGKKDRD*

Z2491 <SEQ ID 3184>

MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSKPAAPVVSEKETEAKEDAPQAGSQG
QGAPSAQGSQDMAAVSEENTGNGGAVTADNPKNEDEVAQNDMPQNAAGTDSSTPNHTPDP
NMLAGNMENQATDAGESSQPANQPDMANAADGMQGDDPSAGGQNAGNTAAQGANQAGNNQ
AAGSSDPIPASNPAPANGGSNFGRVDLANGVLIDGPSQNITLTHCKGDSCSGNNFLDEEV
QLKSEFEKLSDADKISNYKKDGKNDKFVGLVADSVQMKGINQYIFYKPKPTSFARFRRS
ARSRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLTYGAEKLPGGSY
ALRVQGEPAKGEMLAGAAVYNGEVLHFHTENGRPYPTRGRFAAKVDFGSKSVDGIIDSGD
DLHMGTQKFKAAIDGNGFKGTWTENGSGDVSGKFYGPAGEEVAGKYSYRPTDAEKGGFGV
FAGKKEQD*

Figure 21 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 287, further confirming its utility as an antigen for both vaccines and diagnostics.

#### **EXAMPLE 14**

Table 5 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 519 among different strains.

Table 5

519 gene variability: List of used Neisseria strains

**Identification Strains** 

Source / reference

number

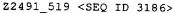
number		
Grou	p B	
zv01_519	NG6/88	R. Moxon / Seiler et al., 1996
zv02_519	BZ198	R. Moxon / Seiler et al., 1996
zv03_519ass	NG3/88	R. Moxon / Seiler et al., 1996
zv04_519	297-0	R. Moxon / Seiler et al., 1996
zv05_519	1000	R. Moxon / Seiler et al., 1996
zv06_519ass	BZ147	R. Moxon / Seiler et al., 1996
zv07_519	BZ169	R. Moxon / Seiler et al., 1996
zv11_519	NGE31	R. Moxon / Seiler et al., 1996
zv12_519	NGF26	R. Moxon / Seiler et al., 1996
zv18_519	BZ232	R. Moxon / Seiler et al., 1996
zv19_519	BZ83	R. Moxon / Seiler et al., 1996
zv20_519ass	44/76	R. Moxon / Seiler et al., 1996
zv21_519ass	MC58	R. Moxon
zv96_519	2996	Our collection
Grouj	ρA	
zv22_519ass	205900	R. Moxon
z2491_519	<b>Z2491</b>	R. Moxon / Maiden et al., 1998
Other		
zv26_519	A22 (group	W) R. Moxon / Maiden et al., 1998
zv27_519	E26 (group	X) R. Moxon / Maiden et al., 1998
zv28_519		oup Y) R. Moxon / Maiden et al., 1998
zv29_519ass		oup Z) R. Moxon / Maiden et al., 1998
Gono	coccus	
zv32_519	Ng F62	R. Moxon / Maiden et al., 1998
fa1090_519	FA1090	R. Moxon

# References:

Seiler A. et al., Mol. Microbiol., 1996, 19(4):841-856. Maiden et al., Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.

The amino acid sequences for each listed strain are as follows:

FA1090_519 <SEQ ID 3185>
MEFFIILLAAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRAMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*



MEFFIILLAAVVVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL ISAGMKIIDSSKTAK*

# ZV01 519 <SEQ ID 3187>

MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL ISAGMKIIDSSKTAK*

#### ZV02 519 <SEQ ID 3188>

MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL ISAGMKIIDSSKTAK*

#### ZV03_519 <SEQ ID 3189>

MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL ISAGMKIIDSSKTAK*

#### ZV04_519 <SEQ ID 3190>

MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL ISAGMKIIDSSKTAK*

# ZV05_519 <SEQ ID 3191>

MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL ISAGMKIIDSSKTAK*

#### ZV06_519ASS <SEQ ID 3192>

MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG RMELDKTFEERDEINSTVFSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERK KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL ISAGMKIIDSSKTAK*

# ZV07_519 <SEQ ID 3193>

MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL ISAGMKIIDSSKTAK*

#### ZV11 519 <SEQ ID 3194>

MEFFIILLAAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL ISAGMKIIDSSKTAK*

#### ZV12 519 <SEQ ID 3195>

MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL ISAGMKIIDSSKTAK*

#### ZV18 519 <SEQ ID 3196>

MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL ISAGMKIIDSSKTAK*

# ZV19 519 <SEQ ID 3197>

MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL ISAGMKIIDSSKTAK*

#### ZV20 519ASS <SEQ ID 3198>

MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSM ISAGMKIIDSSKTAK*

# ZV21_519ASS <SEQ ID 3199>

MEFFTILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL ISAGMKIIDSSKTAK*

#### ZV22_519ASS <SEQ ID 3200>

MEFFIILLAAVVVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE KRARIAESEGRKIEQINLASGQREAKIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL ISAGMKIIDSSKTAK*

## ZV26_519 <SEQ ID 3201>

MEFFIILLAAVVVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL ISAGMKIIDSSKTAK*

ZV27_519 <SEQ ID 3202>

45

MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL ISAGMKIIDSSKTAK*

ZV28_519 <SEQ ID 3203>
MEFFIILLAAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV29_519ASS <SEQ ID 3204>
MEFFIILLAAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSIVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREPEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSNKTAK*

ZV32_519 <SEQ ID 3205>
MEFFIILLAAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRAMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV96_519 <SEQ ID 3206>
MEFFIILLAAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

Figure 22 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 519, further confirming its utility as an antigen for both vaccines and diagnostics.

# **EXAMPLE 15**

Table 6 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 919 among different strains.

Table 6

919 gene variability: List of used Neisseria strains

Identification Strains Source / reference

numb		
	Group B	
zm01	NG6/88	R. Moxon / Seiler et al., 1996
zm02	BZ198	R. Moxon / Seiler et al., 1996
zm03	NG3/88	R. Moxon / Seiler et al., 1996
zm04		R. Moxon / Seiler et al., 1996
<b>zm</b> 05	1000	R. Moxon / Seiler et al., 1996
zm06	BZ147	R. Moxon / Seiler et al., 1996
	BZ169	R. Moxon / Seiler et al., 1996
zm08n	ı <b>52</b> 8	R. Moxon / Seiler et al., 1996
zm09	NGP165	R. Moxon / Seiler et al., 1996
zm10	BZ133	R. Moxon / Seiler et al., 1996
zm11a	sbc NGE31	R. Moxon / Seiler et al., 1996
zm12	NGF26	R. Moxon / Seiler et al., 1996
zm13	NGE28	R. Moxon / Seiler et al., 1996
zm14	NGH38	R. Moxon / Seiler et al., 1996
zm15	SWZ107	R. Moxon / Seiler et al., 1996
zm16	NGH15	R. Moxon / Seiler et al., 1996
<b>zm</b> 17	NGH36	R. Moxon / Seiler et al., 1996
<b>zm</b> 18	BZ232	R. Moxon / Seiler et al., 1996
<b>zm</b> 19	BZ83	R. Moxon / Seiler et al., 1996
zm20	44/76	R. Moxon / Seiler et al., 1996
zm21	MC58	R. Moxon
zm96	2996	Our collection
	Group A	
zm22	205900	R. Moxon
zm23as	sbc F6124	R. Moxon
z2491	Z2491	R. Moxon / Maiden et al., 1998
		, and a second s
	Group C	
zm24	90/18311	R. Moxon
zm25	93/4286	R. Moxon
	Others	
zm26	A22 (group	W) R. Moxon / Maiden et al., 1998
zm27b	(0 1	(group X) R. Moxon / Maiden et al., 1998
	-	oup Y) R. Moxon / Maiden et al., 1998
	₹0-	oup Z) R. Moxon / Maiden et al., 1998
zm31as		amica R. Moxon
		- It is in the second s
	Gonococcus	
	sbc Ng F62	R. Moxon / Maiden et al., 1998
	sbc Ng SN4	R. Moxon
		TO THE PARTY
fa1090	FA1090	R. Moxon
		11. 1.10/1011

# References:

Seiler A. et al., Mol. Microbiol., 1996, 19(4):841-856. Maiden et al., Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.

The amino acid sequences for each listed strain are as follows:

FA1090 <SEQ ID 3207>

MKKHLLRSALYGIAAAILAACQSRSIQTFPQPDTSVINGPDRPAGIPDPAGTTVAGGGAV YTVVPHLSMPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKRFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDGRRTERARFPIYGIPDDFISVPLPAGLRGGKN LVRIRQTGKNSGTIDNAGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSGNEGPVGALGTPLMGEYAGA IDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

Z2491 <SEQ ID 3208>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSVQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL KLGQTSMQGIKAYMQQNPQRLAEVLGQNPSYIFFRELTGSSNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

ZM01 <SEQ ID 3209>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

ZM02 <SEQ ID 3210>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

ZM03 <SEQ ID 3211>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

ZMO4 <SEQ ID 3212>

- A

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVAGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNAGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL KLGQTSMQGIKAYMQQNPQRLAEVLGQNPSYIFFRELTGSSNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

#### ZM05 <SEQ ID 3213>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLSCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

#### ZM06 <SEQ ID 3214>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

#### ZM07 <SEQ ID 3215>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

#### ZMO8N <SEQ ID 3216>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

#### ZM09 <SEQ ID 3217>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVAGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSGNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

#### ZM10 <SEQ ID 3218>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVAGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSGNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

#### ZM11ASBC <SEQ ID 3219>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVGGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSVQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNAGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSSNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

#### ZM12 <SEQ ID 3220>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

#### ZM13 <SEQ ID 3221>

MKKYLFRAALYGIAAAILAACOSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV YTVVPHLSLPHWAEQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

#### ZM14 <SEQ ID 3222>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVAGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNAGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSRNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

# ZM15 <SEQ ID 3223>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDLAGTTVGGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNHQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSGNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

## ZM16 <SEQ ID 3224>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPGRPVGIPDPAGTTVGGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSSNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

ZM17 <SEQ ID 3225>

3.-

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSSNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

#### ZM18 <SEQ ID 3226>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

#### ZM19 <SEQ ID 3227>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

#### ZM20 <SEQ ID 3228>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

# ZM21 <SEQ ID 3229>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

#### ZM22 <SEQ ID 3230>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSVQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL KLGQTSMQGIKAYMQQNPQRLAEVLGQNPSYIFFRELTGSSNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

## ZM23ASBC <SEQ ID 3231>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNAGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL  $\label{thm:confidence} KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA\\ VDRHYITLGAPLFVATAHPVTSKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGETAGK\\ MKEPGYVWQLLPNGMKPEYRP*$ 

#### ZM24 <SEQ ID 3232>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVAGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSGNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

#### ZM25 <SEQ ID 3233>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVAGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGKYMADKGYL KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSGNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

#### ZM26 <SEQ ID 3234>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSVQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL KLGQTSMQGIKAYMQQNPQRLAEVLGQNPSYIFFRELTGSSNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

# ZM27BC <SEQ ID 3235>

MKKYLFRAALYGISAAILAACQSKSIQTFPQPDTSVINGPDRPAGIPDPAGTTVAGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNAGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSSNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGETAGK MKEPGYVWQLLPNGMKPEYRP*

#### ZM28 <SEQ ID 3236>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

## ZM29ASBC <SEQ ID 3237>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL KLGQTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSGNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATTHPITRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

ZM31ASBC <SEQ ID 3238>

MKKHLFRAALYGIAAAILAACQSKSIQTFPQPDTSIIKGPDRPAGIPDPAGTTVGGGGAV YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA LVRIRQTGKNSGTIDNAGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYVFFRELAGSGNDGPVGALGTPLMGEYAGA VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

ZM32ASBC <SEQ ID 3239>

MKKHLLRSALYGIAAAILAACQSRSIQTFPQPDTSVINGPDRPAGIPDPAGTTVAGGGAV YTVVPHLSMPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKRFFER YFTPWQVAGNGSLAGTVTGYYEPVLKGDGRRTERARFPIYGIPDDFISVPLPAGLRGGKA LVRIRQTGKNSGTIDNAGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSGGDGPVGALGTPLMGGYAGA IDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK QKTTGYVWQLLPNGMKPEYRP*

ZM33ASBC <SEQ ID 3240>
MKKHLLRSALYGIAAAILAACQSRSIQTFPQPDTSVINGPDRPAGIPDPAGTTVAGGGAV
YTVVPHLSMPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPIHSFQAKRFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDGRRTERARFPIYGIPDDFISVPLPAGLRGGKN
LVRIRQTGKNSGTIDNAGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMRQNPHKLAEVLGQNPSYIFFRELAGSGNEGPVGALGTPLMGEYAGA
IDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM96 <SEQ ID 3241>
MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIQGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

Figure 23 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 919, further confirming its utility as an antigen for both vaccines and diagnostics.

# **EXAMPLE 16**

Using the above-described procedures, the following oligonucleotide primers were employed in the polymerase chain reaction (PCR) assay in order to clone the ORFs as indicated:

Table 7: Oligonucleotides used for PCR to amplify complete or partial ORFs

001			
1.001			Restriction sites
001	Forward	CGCGGATCCCATATG-TGGATGGTGCTGGTCAT	BamHI-
	Darrana	CCCCCTCC A C TCCCCTCTTCTTCTCTCTCTCTCTCTC	NdeI
003	Reverse	CCCGCTCGAG-TGCCGTCTTGTCCCAC	XhoI
003	Forward	CGCGGATCCCATATG-GTCGTATTCGTGGC	BamHI-
	Reverse	CCCGCTCGAG-AAAATCATGAACACGCGC	NdeI
005	Forward	CGCGGATCCCATATC CAGAATATTTCAGAGATATT	XhoI
005	1 Of Ward	CGCGGATCCCATATG-GACAATATTGACATGT	BamHI-
	Reverse	CCCGCTCGAG-CATCACATCCGCCCG	NdeI
006	Forward	CGCGGATCCCATATG-CTGCTGGTGCTGG	XhoI
	2 02 11 02 0	esessing cide ide ide ide ide	BamHI-
	Reverse	CCCGCTCGAG-AGTTCCGGCTTTGATGT	NdeI
007	Forward	CGCGGATCCCATATG-GCCGACAACAGCATCAT	XhoI
		COOGNOTMONOCATORI	BamHI- NdeI
	Reverse	CCCGCTCGAG-AAGGCGTTCATGATATAAG	XhoI
800	Forward	CGCGGATCCCATATG-AACAACAGACATTTTG	BamHI-
			NdeI
	Reverse	CCCG <u>CTCGAG</u> -CCTGTCCGGTAAAAGAC	XhoI
009	Forward	CGCGGATCCCATATG-CCCCGCGCTGCT	BamHI-
	D	GGGGGMGG LG The service	NdeI
011	Keverse	CCCGCTCGAG-TGGCTTTTGCCACGTTTT	XhoI
011	rorward (	CGCGGATCCCATATG-AAGACACCCGCAAG	BamHI-
	Reverse (	CCCCCTCCAC CCCCCTCACTACCT	NdeI
		CCCGCTCGAG-GGCGGTCAGTACGGT	XhoI
012	roiwaid v	CGCGGATCCCATATG-CTCGCCCGTTGCC	BamHI-
	Reverse (	CCCG <u>CTCGAG</u> -AGCGGGGAAGAGGCAC	NdeI
013	Forward (	CGCGGATCCCATATG-CCTTTGACCATGCT	XhoI
		-cerridaccarder	BamHI-
	Reverse (	CCCG <u>CTCGAG</u> -CTGATTCGGCAAAAAATCT	NdeI
018	Forward (	CGCGGATCCCATATG-CAGCAGAGGCAGTT	XhoI
			BamHI- NdeI
	Reverse (	CCCGCTCGAG-GACGAGGCGAACGCC	XhoI
		AAAGAATTC-CTGCCAGCCGGCAAGACCCCGGC	Eco RI
	Reverse A	AAACTGCAG-TCAGCGGGCGGGGACAATGCCCAT	Pst I
023	Forward A	AAAGAATTC-AAAGAATATTCGGCATGGCAGGC	Eco RI
	Reverse A	AAACTGCAG-TTACCCCCAAATCACTTTAACTGA	Pst I
025	Forward A	AAAGAATTC-TGCGCCACCCAACAGCCTGCTCC	Eco RI
	Reverse A	AAACTGCAG-TCAGAACGCGATATAGCTGTTCGG	Pst I
031 1	Forward C	CGCGGATCCCATATG-GTCTCCCTTCGCTT	BamHI-
			Ndel
]	Reverse C	CCGCTCGAG-ATGTAAGACGGGGACAAC	XhoI
032 I	Forward C	GCGGATCCCATATG-CGGCGAAACGTGC	BamHI-

<u></u>			
	<b>D</b>	COCCOTOCA C. CTC CTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	Ndel
022		CCCGCTCGAG-CTGGTTTTTTGATATTTGTG	XhoI
033	Forward	CGCGGATCCCATATG-GCGGCGGCAGACA	BamHI-
1	D	COCCOTOCA O ATTROCCO A TOCCOTO	NdeI
024		CCCGCTCGAG-ATTTGCCGCATCCCGAT	XhoI
034	Forward	CGCGGATCCCATATG-GCCGAAAACAGCTACGG	BamHI-
	D	CCCCCTCC A C TTTC A CC A TTTC C CTTC C CTTC	NdeI
026		CCCGCTCGAG-TTTGACGATTTGGTTCAATT	XhoI
036	Forward	CGCGGATCCCATATG-CTGAAGCCGTGCG	BamHI-
	Davierse	CCCCCTCCAC CCCCACTCCCTATACCC	NdeI
020		CCCGCTCGAG-CCGGACTGCGTATCGG	XhoI
038	Forward	CGCGGATCCCATATG-ACCGATTTCCGCCA	BamHI-
	Daviere	CCCCCTCC A C TTCT A CCCCCT A CTCCC	NdeI
020		CCCGCTCGAG-TTCTACGCCGTACTGCC	Xhol
039	Forward	CGCGGATCCCATATG-CCGTCCGAACCGC	BamHI-
1	Dovorco	CCCGCTCGAG-TAGGATGACGAGGTAGG	NdeI
041	Forward	CCCCCATCCCATATC TTCCTCCCCCA + CCCC	XhoI
041	roiwaid	CGCGGATCCCATATG-TTCGTGCGCGAACCGC	BamHI-
	Deverse	CCCGCTCGAG-GCCCAAAAACTCTTTCAAA	NdeI
042			XhoI
042	roiwatu	CGCGGATCCCATATG-ACGATGATTTGCTTGC	BamHI-
	Reverse	CCCGCTCGAG-TTTGCAGCCTGCATTTGAC	NdeI
043		AAAAAAGGTACC-ATGGTTGTTTCAAATCAAAATATC	XhoI
043		AAACTGCAG-TTATTGCGCTTCACCTTCCGCCGC	Kpn I
0432	Forward	AAAAAAGGTACC GCAAAACTCGATGGGGGTTGGAGGGTTGGA	Pst I
043a	Deverse	AAAAAAGGTACC-GCAAAAGTGCATGGCGGCTTGGACGGTGCAAAAAACTGCAG-	1
	Keveise	TTAATCCTGCAACACGAATTCGCCCGTCCG	Pst I
044	Forward	CGCGGATCCCATATG-CCGTCCGACTAGAG	
• • •	1 OI Wald	ede <u>dentecentard</u> -ecorecoactadad	BamHI-
	Reverse	CCCGCTCGAG-ATGCGCTACGGTAGCCA	NdeI
046		AAAGAATTC-ATGTCGGCAATGCTCCCGACAAG	XhoI Eco RI
		AAACTGCAG-TCACTCGGCGACCCACACCGTGAA	1
047	Forward	CGCGGATCCCATATG-GTCATCATACAGGCG	Pst I
• · ·	1 01 11 11	ede <u>dentedentitio</u> -dicarcaracadded	BamHI-
•	Reverse	CCCGCTCGAG-TCCGAAAAAGCCCATTTTG	NdeI XhoI
048		AAAGAATTC-ATGCTCAACAAAGGCGAAGAATTGCC	
		AAACTGCAG-TCAAGATTCGACGGGGATGATGCC	Eco RI
049		AAAGAATTC-ATGCGGGCGCAGGCGTTTGATCAGCC	Pst I
0.5		AAACTGCAG-AAGGCGTATCTGAAAAAATGGCAG	Eco RI
050		CGCGGATCCCATATG-GGCGCGGGCTGG	Pst I
	1 Ol Wala	ede <u>ddareeeararu</u> -ddededddc1dd	BamHI-
	Reverse	CCCGCTCGAG-AATCGGGCCATCTTCGA	NdeI
052		AAAAAAGAATTC-ATGGCTTTGGTGGCGGAGGAAAC	XhoI
		AAAAAAGTCGAC-TCAGGCGGCGTTTTTCACCTTCCT	Eco RI
052a	Forward	A A A A A C A ATTC CTCCCCCA CCA A A CCC	Sal I
		GIGGGGAAAACGGAAAIAICCGC	Eco RI

	Reverse AAAAAACTGCAG-TTAGCTGTTTTTTGGAAACGCCGTCCAACC	
073	Forward CGCGGATCCCATATG-TGTATGCCATATAAGAT	
10,5	Tolward ededdateceataid-idiaidecatataagai	BamHI-
	Reverse CCCGCTCGAG-CACCGGATTGTCCGAC	NdeI
075		XhoI
0,3	Tolward CocodATCCCATATO-CCGTCTTACTTCATC	BamHI-
l	Reverse CCCGCTCGAG-ATCACCAATGCCGATTATTT	Ndel
0772	Forward AAAAAAGAATTC-GGCGGCATTTTCATCGACACCTTCCT	XhoI
""	Reverse AAAAACTGCAG TCACAACAA GATGTGGAA GATGTGGAAGATTCTCT	Eco RI
080	Reverse AAAAAACTGCAG-TCAGACGAACATCTGCACAAACGCAAT	Pst I
000	Torward TEERIGIETTE GOOT COOLOGITTING TITTACAATTO	Eco RI
081	Reverse AAACTGCAG-CTATTCTTCGGGATTCTTTTCGGG	Pst I
001	Forward AAAGAATTC-ATGAAACCACTGGACCTAAATTTCATCTG	Eco RI
082	Reverse AAACTGCAG-TCACTTATCCTCCAATGCCTC	Pst I
002	Forward AAAGAATTC-ATGTGGTTGTTGAAGTTGCCTGC	Eco RI
004	Reverse AAACTGCAG-TTACGCGGATTCGGCAGTTGG	Pst I
084	Forward AAAGAATTC-TATCACCCAGAATATGAATACGGCTACCG	Eco RI
005	Reverse AAACTGCAG-TTATACTTGGGCGCAACATGA	Pst I
085	Forward CGCGGATCCCATATG-GGTAAAGGGCAGGACT	BamHI-
İ	Payrama CCCCCTCCAC CAAACCCTTA	NdeI
086	Reverse CCCGCTCGAG-CAAAGCCTTAAACGCTTCG	XhoI
080	TO THE THE HELD THE THE THE THE THE THE THE THE THE THE	Kpn I
007	Reverse AAACTGCAG-TTACTCCACCGATAACCGCG	Pst I
067	Forward AAAGAATTC-ATGGGCGGTAAAACCTTTATGC	Eco RI
007-	Reverse AAACTGCAG-TTACGCCGCACACGCAATCGC	Pst I
08/a	Forward AAAAAAGAATTC-AAGCTATTAGGCGTGCCGATTGTGATTCA	Eco RI
000	Reverse AAAAAACTGCAG-TTACGCCTGCAAGATGCCCAGCTTGCC	Pst I
UOO	Forward AAAAAAGAATTC-ATGTTTTTATGGCTCGCACATTTCAG	Eco RI
000	Reverse AAAAAACTGCAG-TCAGCGGATTTTGAGGGTACTCAAACC	Pst I
089	Forward CGCGGATCCCATATG-CCGCCCAAAATCAC	BamHI-
	Reverse CCCGCTCGAG-TGCGCATACCAAAGCCA	NdeI
090	Forward CGCGGATCCCATATC CGCATA GTGG A GG	XhoI
070	Forward CGCGGATCCCATATG-CGCATAGTCGAGCA	BamHI-
	Reverse CCCGCTCGAG-AGCAAAACGGCGGTACG	NdeI
091	Forward AAAGAATTC-ATGGAAATACCCGTACCGCCGAGTCC	XhoI
	Reverse AAACTGCAG-TCAGCGCAGGGGGTAGCCCAAGCC	Eco RI
092	Forward AAAGAATTC-ATGTTTTTTTTTTTCAATCCG	Pst I
~~ <b>~</b>	Reverse AAACTGCAG-TCAAATCTGTTTCGACAATGC	Eco RI
093	Forward AAAGAATTC-ATGCAGAATTTTGGCAAAGTGGC	Pst I
0,5	Reverse AAACTGCAG-CTATGGCTCATACCGGGC	Eco RI
094	Forward AAAGAATTC ATCCCCTCAGGGAAGGGGC	Pst I
J) T	Forward AAAGAATTC-ATGCCGTCACGGAAGCGCATCAACTC Reverse AAACTGCAG-TTATCCCGGCCATACCGCCGAACA	Eco RI
095	Forward AAAGAATTC ATGTCGTTTGATTTTGATTTTGATTTTGATTTTGATTTTGATTTTGATTTTGATTTTGATTTTGATTTTGATTTTGATTTTGATTTTGATTTTGATTTTTGATTTTTT	Pst I
	Forward AAAGAATTC-ATGTCCTTTCATTTGAACATGGACGG  Reverse AAACTGCAG TCAACCGCCAACGGACTA	Eco RI
096	Reverse AAACTGCAG-TCAACGCCGCAGGCACTAACGCCC Forward AAAGAATTC ATGCCTCATCATAGGCCCC	Pst I
	Forward AAAGAATTC-ATGGCTCGTCATACCGGGCAGGG	Eco RI

	Doverse	A A A CTGC A G TC A A A CC A A A A GGGGGGTGTG	
097	Famusand	AAACTGCAG-TCAAAGGAAAAGGCCGTCTGAAAAGCG	Pst I
097		AAAGAATTC-ATGGACACTTCAAAACAAACACTGTTG	Eco RI
		AAACTGCAG-TCAGCCCAAATACCAGAATTTCAG	Pst I
098	Forward	AAAGAATTC-GATGAACGCAGCCCAGCATGGATACG	Eco RI
		AAACTGCAG-TTACGACATTCTGATTTGGCA	Pst I
102	Forward	AAAAAAGAATTC-GGCCTGATGATTTTGGAAGTCAACAC	Eco RI
		AAAAAACTGCAG-TTATCCTTTAAATACGGGGACGAGTTC	Pst I
105	Forward	CGCGGATCCCATATG-TCCGCAAACGAATACG	BamHI-
İ			NdeI
		CCCGCTCGAG-GTGTTCTGCCAGTTTCAG	XhoI
107	Forward	AAAAAGAATTC-	Eco RI
	_	CTGATGATTTTGGAAGTCAACACCCATTATCC	
	Reverse	AAAAAACTGCAG-TTATCCTTTAAATACGGGGACGAGTTC	Pst I
[1076	Forward	AAAAAGAATTC-	Eco RI
	D	GATACCCAAGCCCCCGCCGGCACAAACTACTG	
	Reverse	AAAAAACTGCAG-	Pst I
108	Forward	TTACGCGTCGCCTTTAAAGTATTTGAGCAGGCTGGAGAC AAAGAATTC-ATGTTGCCGGGCTTCAACCG	_
100		AAACTGCAG-TTAGCGGTACAGGTGTTTGAAGCA	Eco RI
1082			Pst I
1002	Pavarce	AAAAAAGAATTC-GGTAACACATTCGGCAGCTTAGACGGTGGAAACTGCAG-TTAGCGGTACAGGTGTTTGAAGCA	
100	Forward	AAAGAATTC-ATGTATTATCGCCGGGTTATGGG	Pst I
103			Eco RI
111		AAACTGCAG-CTAGCCCAAAGATTTGAAGTGTTC	Pst I
111	roiwaiu	CGCGGATCCCATATG-TGTTCGGAACAAACCGC	BamHI-
	Reverse	CCCGCTCGAG-GCGGAGCAGTTTTTCAAA	NdeI
114		CGCGGATCCCATATG-GCTTCCATCACTTCGC	XhoI
** '	1 OI Wala	ede <u>ddrifeceataid</u> -deffecateaeffede	BamHÏ⇔
	Reverse	CCCGCTCGAG-CATCCGCGAAATCGTC	NdeI XhoI
117		AAAAAAGGTACC-ATGGTCGAAGAACTGCTG	
		AAACTGCAG-TTAAAGCCGGGTAACGCTCAATAC	Kpn I
118		AAAGTCGACATGTGTGAGTTCAAGGATATTATAAG	Pst I
ļ	Reverse	AAAGCATGC-CTATTTTTGTTGTAATAATCAAATC	Sal I
121		CGCGGATCCCATATG-GAAACACAGCTTTACAT	Sph I
		GRANCHORGETTROAT	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATAATAATATCCCGCGCCC	XhoI
122		CGCGGATCCCATATG-GTCATGATTAAAATCCGCA	BamHI-
			Ndel
		CCCGCTCGAG-AATCTTGGTAGATTGGATTT	XhoI
125		AAAGAATTC-ATGTCGGGCAATGCCTCCTCTCC	Eco RI
	Reverse	AAACTGCAG-TCACGCCGTTTCAAGACG	Pet I
125a	Forward .	AAAAAAGAATTC-ACGGCAGGCAGCACCGCCGCACAGGTTTC	Eco RI
	Reverse .	AAAAAA <u>CTGCA</u> G-	Pst I
		TTATTTTGCCACGTCGGTTTCTCCGGTGAACAACGC	- 50.1
126	Forward	CGCGGATCCCATATG-CCGTCTGAAACCC	BamHI-

1	Reverse CCCGCTCGAG-ATATTCCGCCGAATGCC	Ndel
12	7 Forward AAAGAATTC ATCCAAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCCAATTATCAATTATCCAATTATCAATTATCCAATTATCCAATTATCAATTATCCAATTATCAATTATCCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATCAATTATC	XhoI
	Forward AAAGAATTC-ATGGAAATATGGAATATGTTGGACACTTG Reverse AAACTGCAG-TTAAAGTGTTTCGGAGCCGGC	Eco RI
127	a Forward AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	Pst I
12,	a Forward AAAAAAGAATTC-AAGGAACTGATTATGTGTCTGTCGGG	Eco RI
128	Reverse AAACTGCAG-TTAAAGTGTTTCGGAGCCGGC	Pst I
120	Forward CGCGGATCCCATATG-ACTGACAACGCACT	BamHI-
	Reverse CCCGCTCGAG-GACCGCGTTGTCGAAA	NdeI
130	Forward CGCGGATCCCATATG-AAACAACTCCGCGA	XhoI
	TOTALIA COCUMENTO ANACAACTICOGOGA	BamHI-
	Reverse CCCGCTCGAG-GAATTTTGCACCGGATTG	NdeI
132		XhoI
	Reverse AAAAAACTGCAG-TCACCATGTCGGCATTTGAAAAAC	Eco RI
134	Forward CGCGGATCCCATATG-TCCCAAGAAATCCTC	Pst I
	- TO - TO COLANDA AT COLO	BamHI-
	Reverse CCCGCTCGAG-CAGTTTGACCGAATGTTC	Ndel
135	Forward CGCGGATCCCATATG-AAATACAAAAGAATCGTATT	XhoI BamHI-
		NdeI
	Reverse CCCGCTCGAG-AAATTCGGTCAGAAGCAGG	XhoI
137	TOTAL TERRETORIES AT TACCCATCCCAAT TCGATCC	Kpn I
	Reverse AAAAAACTGCAG-TCAGTGCTGTTTTTTCATGCCGAA	Pst I
137a	Forward AAAAAAGAATTC-GGCCGCAAACACGGCATCGGCTTCCT	Eco RI
	Reverse AAAAAACTGCAG-TTAAGCGGGATGACGCGGCAGCATACC	Pst I
138	Forward AAAAAAGAATTC-AACTCAGGCGAAGGAGTGCTTGTGGC	Eco RI
	Reverse AAAAAATCTAGA-TCAGTTTAGGGATAGCAGGCGTAC	Xba I
141	Forward AAAGAATTC-ATGAGCTTCAAAACCGATGCCGAAATCGC	Eco RI
	Reverse AAACTGCAG-TCAGAACAAGCCGTGAATCACGCC	Pst I
142	Forward CGCGGATCCCATATG-CGTGCCGATTTCATG	BamHI-
		NdeI
	Reverse CCCGCTCGAG-AAACTGCTGCACATGGG	XhoI
143	Forward AAAAAGAATTC-	Eco RI
	ATGCTCAGTTTCGGCCTTCAGAC	
1 4 4	Reverse AAAAAACTGCAG-TCAAACCCCGCCGTGTGTTTCTTTAAT	Pst I
144	Forward AAAAAAGAATTC-GGTCTGATCGACGGGCGTGCCGTAAC	Eco RI
146	Reverse AAAAAATCTAGA-TCGGCATCGGCCGGCATATGTCCG	Xba I
140	Forward AAAAAGAATTC-	Eco RI
	CGCCAAGTCGTCATTGACCACGACAAAGTC	
147	Reverse AAAAAACTGCAG-TTAGGCATCGGCAAATAGGAAACTGGG	Pst I
1 <b>7</b> /	Forward AAAAAAGAATTC-ACTGAGCAATCGGTGGATTTGGAAAC	Eco RI
148	Reverse AAAAAATCTAGA-TTAGGTAAAGCTGCGGCCCATTTGCGG Forward AAAAAAGAATTC-	Xba I
140		Eco RI
	ATGGCGTTAAAAACATCAAACTTGGAACACGC Reverse AAAAAATCTAGA-TCAGCCCTTCATACAGCCTTCGTTTTG	
149	Forward CGCGGATCCCATATG-CTGCTTGACAACAAAGT	Xba I
		BamHI-

			<del>-                                    </del>
	Deverse	CCCGCTCGAG-AAACTTCACGTTCACGCC	NdeI
150			XhoI
130	rorward	CGCGGATCCCATATG-CAGAACACAAATCCG	BamHI-
ĺ	Daverce	CCCGCTCGAG-ATAAACATCACGCTGATAGC	NdeI
151		AAAAAGAATTC-	XhoI
131	rorwaru	ATGAAACAATCCGCAACATCGCCATCATCGC	Eco RI
	Daverce	AAAAAACTGCAG-TCAATCGCCATCATCGC	***
152		AAAAAGAATTC-	Pst I
132	roiwaiu	ATGAAAAACAAAACCAAAGTCTGGGACCTCCC	Eco RI
	Reverse	AAAAACTGCAG-TCAGGACAGGAGCAGGATGGCGGC	D . T
153		AAAAAGAATTC-ATGGCGTTTGCTTACGGTATGAC	Pst I
133		AAAAACTGCAG-TCAGTCATGTTTTTCCGTTTCATT	Eco RI
1520			Pst I
133a		AAAAAAGAATTC-CGGACTTCGGTATCGGTTCCCCAGCATTG AAAAAACTGCAG-	Eco RI
	Reverse	TTACGCCGACGAAATACTCAGACTTTTCGG	Pst I
154	Forward	CGCGGATCCCATATG-ACTGACAACAGCCC	
154	Torward	COCOGATECEATATO-ACTOACAACAGCCC	BamHI-
}	Reverse	CCCGCTCGAG-TCGGCTTCCTTTCGGG	NdeI
155		AAAAAGAATTC-ATGAAAATCGGTATCCCACGCGAGTC	XhoI
133		AAAAAACTGCAG-TTACCCTTTCTTAAACATATTCAGCAT	Eco RI
156		AAAAAAGAATTC-GCACAGCAAAACGGTTTTGAAGC	Pst I
130		AAAAACTGCAG-TCAAGCAGCCGCGACAAACAGCCC	Eco RI
157			Pst I
13/	rorward	CGC <u>GGATCCCATATG</u> -AGGAACGAGGAAAAAC	BamHI-
	Reverse	CCCGCTCGAG-AAAACACAATATCCCCGC	NdeI
158		AAAAAAGAATTC-GCGGAGCAGTTGGCGATGGCAAATTCTGC	XhoI
150		AAAAATCTAGA-TTATCCACAGAGATTGTTTCCCAGTTC	. ***
160		CGCGGATCCCATATG-GACATTCTGGACAAAC	Xba I
100	TOIWalu	COCOUNTECCATATO-OACATICIOGACAAAC	BamHI-
	Reverse	CCCGCTCGAG-TTTTTGCCCGCCTTCTTT	NdeI
163		AAAAAAGGTACC-ACCGTGCCGGATCAGGTGCAGATGTG	XhoI
103		AAAAATCTAGA-TTACTCTGCCAATTCCACCTGCTCGTG	Kpn I
1639		AAAAAAGAATTC-CGGCTGGTGCAGATAATGAGCCAGAC	Xba I
1054		AAAAATCTAGA-TTACTCTGCCAATTCCACCTGCTCGTG	Eco RI
164			Xba I
104	roiward	CGCGGATCCCATATG-AACCGGACTTATGCC	BamHI-
	Reverse	CCCGCTCGAG-TTTGTTTCCGTCAAACTGC	NdeI
165		CGCGGATCCGCTAGC-GCTGAAGCGACAGACG	XhoI
103	TOIWALU	COCOCATCCOCTAGC-OCTGAAGCGACAGACG	BamHI-
	Reverse	CCCGCTCGAG-AATATCCAATACTTTCGCG	NheI XhoI
206		CGCGGATCCCATATG-AAACACCGCCAACCGA	
	- OI WALL	OOO OO MAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	BamHI-
	Reverse	CCCGCTCGAG-TTCTGTAAAAAAAGTATGTGC	NdeI
209		CGCGGATCCCATATG-CTGCGGCATTTAGGA	XhoI
		<u></u> orgodominadox	BamHI- NdeI
			TAGET

	Daviero CCCCCTCC AC THE CCCCTC	
211	Reverse CCCGCTCGAG-TACCCCTGAAGGCAAC	XhoI
211	Toward In the Factor of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control of the Control	Eco RI
	Reverse AAAAAACTGCAG-CTATCCTGCGGATTGGCATTGAAA	Pst I
212	Forward CGCGGATCCCATATG-GACAATCTCGTATGG	BamHI-
-		NdeI
1	Reverse CCCGCTCGAG-AGGGGTTAGATCCTTCC	XhoI
215	Forward CGCGGATCCCATATG-GCATGGTTGGGTCGT	BamHI-
	D	NdeI
	Reverse CCCGCTCGAG-CATATCTTTTGTATCATAAATC	XhoI
216	Forward CGCGGATCCCATATG-GCAATGGCAGAAAACG	BamHI-
	D 000000000000000000000000000000000000	NdeI
	Reverse CCCGCTCGAG-TACAATCCGTGCCGCC	XhoI
217	Forward CGCGGATCCCATATG-GCGGATGACGGTGTG	BamHI-
	D CCCCCTCC C CCCC	NdeI
210	Reverse CCCGCTCGAG-ACCCCGAATATCGAATCC	XhoI
218	Forward CGCGGATCCCATATG-GTCGCGGTCGATC	BamHI-
-	D	NdeI
210	Reverse CCCGCTCGAG-TAACTCATAGAATCCTGC	$\mathbf{X}$ ho $\mathbf{I}$
219	Forward CGCGGATCCGCTAGC-ACGGCAAGGTTAAG	BamHI-
	Parione CCCCCTCCAC TTTA A A CC A TOTA	NheI
222	Reverse CCCGCTCGAG-TTTAAACCATCTCCTCAAAAC	XhoI
223	Forward CGCGGATCCCATATG-GAATTCAGGCACCAAGTA	BamHI-
	Reverse CCCGCTCGAG CCCTTCGGGGGGTGTG	NdeI
225	Reverse CCCGCTCGAG-GGCTTCCCGCGTGTC	$\mathbf{X}$ ho $\mathbf{I}$
223	Forward CGCGGATCCCATATG-GACGAGTTGACCAACC	BamHI-
	Reverse CCCGCTCGAG-GTTCAGAAAGCGGGAC	NdeI
226	Forward AAAGAATTC-CTTGCGATTATCGTGCGCACGCG	XhoI
220	Peverse AAACTCCAC TCAAAATCCGAAAACCGCG	Eco RI
228	Reverse AAACTGCAG-TCAAAATCCCAAAACGGGGAT	Pst I
220	Forward CGCGGATCCCATATG-TCGCAAGAAGCCAAACAG	BamHI-
	Reverse CCCGCTCGAG-TTTGGCGGCATCTTTCAT	Ndel
229	Forward CGCGGATCCCATATC CAACACCTTTTTCGC	XhoI
22)	Forward CGCGGATCCCATATG-CAAGAGGTTTTGCCC	BamHI-
	Reverse CCCGCTCGAG-ACACAATATAGCGGATGAAC	NdeI
230	Forward CGCGGATCCCATATG-CATCCGGGTGCCGAC	XhoI
250	Tolwan Cocada CCCATATO-CATCCGGGTGCCGAC	BamHI-
	Reverse CCCGCTCGAG-AAGTTTGGCGGCTTCGG	NdeI
232	Forward AAAAAAGAATTC-ATGTACGCTAAAAAAGGCGGTTTGGG	XhoI
	Reverse AAAAAACTGCAG-TCAAGGTTTTTCCTGATTGCCGCCGC	Eco RI
232a	Forward AAAAAAAAAAAATTC CCCAACCCTCCCCAATTCCCCCCCC	Pst I
<b></b>	Forward AAAAAAGAATTC-GCCAAGGCTGCCGATACACAAATTGA	Eco RI
233	Reverse AAAAAACTGCAG-TTAAACATTGTCGTTGCCGCCCAGATG	Pst I
433	Forward CGCGGATCCCATATG-GCGGACAAACCCAAG	BamHI-
	Reverse CCCGCTCGAG GACCGCATTTGAGGGG	NdeI
234	Reverse CCCGCTCGAG-GACGGCATTGAGCAG  Forward CGCGGATCGCATATG CGCGTTTTGAGCAG	XhoI
4J7	Forward CGCGGATCCCATATG-GCCGTTTCACTGACCG	BamHI-
		<del></del>

	Derroma	CCCCA A CCTT A CCCTTCC A TTTCCCA TTC	Ndel
225		GCCCAAGCTT-ACGGTTGGATTGCCATG	Hind III
235	Forward	CGCGGATCCCATATG-GCCTGCCAAGTTCAAA	BamHI-
	Darramaa	CCCCCTCC & C TTTCCCCCTCCTTCCTTCTTCC	NdeI
226		CCCGCTCGAG-TTTGGGCTGCTCTTC	XhoI
236	Forward	CGCGGATCCCATATG-GCGCGTTTCGCCTT	BamHI-
	D	CCCCCTCC A C ATCCCTCCCCCCCC	NdeI
220		CCCGCTCGAG-ATGGGTCGCGCGCCGT	XhoI
238	rorward	CGCGGATCCGCTAGC-AACGGTTTGGATGCCCG	BamHI-
ļ	Darramas		NheI
220		CCCGCTCGAG-TTTGTCTAAGTTCCTGATATG	XhoI
239	Forward	CCGGAATTCTACATATG-CTCCACCATAAAGGTATTG	EcoRI-
	Dorrora	CCCGCTCGAG-TGGTGAAGAGCGGTTTAG	NdeI
240			XhoI
240	Forward	CGCGGATCCCATATG-GACGTTGGACGATTTC	BamHI-
	Deverse	CCCGCTCGAG-AAACGCCATTACCCGATG	NdeI
241			XhoI
241	Polwaru	CCGGAATTCTACATATG-CCAACACGTCCAACT	EcoRI-
	Reverse	CCCGCTCGAG-GAATGCGCCTGTAATTAATC	Ndel
242		CGCGGATCCCATATG-ATCGGCAAACTTGTTG	XhoI
272	Torward	COCOCATATO-ATCOCCAAACTIGITG	BamHI-
	Reverse	GCCCAAGCTT-ACCGATACGGTCGCAG	NdeI
243		CGCGGATCCCATATG-ACGATTTTTTCGATGCTGC	HindIII
273	1 OI Wald	ede <u>ddrifeeeriri</u> g-redrifffffedrigeIge	BamHI-
	Reverse	CCCGCTCGAG-CGACTTGGTTACCGCG	NdeI
244		CGCGGATCCCATATG-CCGTCTGAAGCCC	XhoI
		control control	BamHI-
	Reverse	CCCGCTCGAG-TTTTTTCGGTAGGGGATTT	NdeI* XhoI
246		CGCGGATCCCATATG-GACATCGGCAGTGC	BamHI-
		<u> </u>	Ndel
	Reverse	CCCGCTCGAG-CCCGCGCTGCTGGAG	XhoI
247		CGCGGATCCCATATG-GTCGGATCGAGTTAC	BamHI-
			Ndel
	Reverse	CCCGCTCGAG-AAGTGTTCTGTTTGCGCA	XhoI
248	Forward	CGCGGATCCCATATG-CGCAAACAGAACACT	BamHI-
			Ndel
	Reverse	CCCGCTCGAG-CTCATCATTATTGCTAACA	XhoI
249	Forward	CGCGGATCCCATATG-AAGAATAATGATTGCTTC	BamHI-
			NdeI
		CCCGCTCGAG-TTCCCGACCTCCGAC	XhoI
251	Forward	CGCGGATCCCATATG-CGTGCTGCGGTAGT	BamHI-
		· ———	Ndel
		CCCG <u>CTCGAG</u> -TACGAAAGCCGGTCGTG	XhoI
253	Forward	AAAAAAGAATTC-ATGATTGACAGGAACCGTATGCTGCG	Eco RI
	Reverse	AAAAAACTGCAG-TTATTGGTCTTTCAAACGCCCTTCCTG	Pst I

253	Forward AAAAAGAATTC AAAATCCTTTTTCAAAACCA	
2336	Forward AAAAAAGAATTC-AAAATCCTTTTGAAAACAAGCGAAAACG	
251	Reverse AAAAAACTGCAG-TTATTGGTCTTTCAAACGCCCTTCCTG	Pst I
254	Forward AAAAAAGAATTC-ATGTATACAGGCGAACGCTTCAATAC	Eco RI
255	Reverse AAAAAATCTAGA-TCAGATTACGTAACCGTACACGCTGAC	Xba I
233	Forward CGCGGATCCCATATG-GCCGCGTTGCGTTAC	BamHI-
	Pewerse CCCCCTCCAC ATCCCCA ATLACCA ATLACCA	NdeI
256	Reverse CCCGCTCGAG-ATCCGCAATACCGACCAG	XhoI
230	Forward CGCGGATCCGCTAGC-TTTTAACACCGCCGGAC	BamHI-
	Reverse CCCGCTCGAG-ACGCCTGTTTGTGCGG	NheI
257		XhoI
25,	Toward ede <u>ddateeeatatd</u> -deddiffefffeff	BamHI-
	Reverse CCCGCTCGAG-GCGCGTGAATATCGCG	NdeI
258	Forward AAAAAAGAATTC-GATTATTTCTGGTGGATTGTTGCGTTCAG	XhoI
	Reverse AAAAAACTGCAG-CTACGCATAAGTTTTTACCGTTTTTGG	Eco RI
258a	Forward AAAAAAGAATTC-GCGAAGGCGGTGGCGCAAGGCGA	Pst I
	Reverse AAAAAACTGCAG-CTACGCATAAGTTTTTACCGTTTTTGG	Eco RI
259	Forward CGCGGATCCCATATG-GAAGAGCTGCCTCCG	Pst I
	TOTAL COCCONTICCCATATO-GAAGAGCTGCCTCCG	BamHI-
	Reverse CCCGCTCGAG-GGCTTTTCCGGCGTTT	NdeI
<b>26</b> 0	Forward CGCGGATCCCATATG-GGTGCGGGTATGGT	XhoI
	- 10 OST COST TITO OG TO COO TA TOG T	BamHI-
	Reverse CCCGCTCGAG-AACAGGGCGACACCCT	NdeI XhoI
261		Eco RI
	Reverse AAAAAACTGCAG-TCAAACCAACAAGCCTTGGTCACT	
263	Forward CGCGGATCCCATATG-GCACGTTTAACCGTA	Pst I BamHI-
		Ndel
	Reverse CCCGCTCGAG-GGCGTAAGCCTGCAATT	XhoI
264	Forward AAAAAAGGTACC-GCCGACGCAGTGGTCAAGGCAGAA	Kpn I
	Reverse AAACTGCAG-TCAGCCGGCGGTCAATACCGCCCG	Pst I
265	Forward AAAAAAGAATTC-GCGGAGGTCAAGAGAAGGTGTTTG	Eco RI
	Reverse AAAAAACTGCAG-TTACGAATACGTCGTCAAAATGGG	Pst I
266	Forward AAAGAATTC-CTCATCTTTGCCAACGCCCCCTTC	Eco RI
	Reverse AAACTGCAG-CTATTCCCTGTTGCGCGTGTGCCA	Pst I
267	Forward AAAGAATTC-TTCTTCCGATTCGATGTTAATCG	Eco RI
	Reverse AAACTGCAG-TTAGTAAAAACCTTTCTGCTTGGC	Pst I
269	Forward AAAGAATTC-TGCAAACCTTGCGCCACGTGCCC	Eco RI
	Reverse AAACTGCAG-TTACGAAGACCGCAACGAAAGGCAGAG	Pst I
269a	Forward AAAAAAGAATTC-GACTTTATCCAAAACACGGCTTCGCC	Eco RI
	Reverse AAACTGCAG-TTACGAAGACCGCAACGAAAGGCAGAG	
270	Forward AAAGAATTC-GCCGTCAAGCTCGTTTTGTTGCAATG	Pst I
	Reverse AAACTGCAG-TTATTCGGCGGTAAATGCCGTCTG	Eco RI
271	Forward CGCGGATCCCATATG-CCTGTGTGCAGCTCGAC	Pst I
	COTOTOTOCACTCOAC	BamHI-
	Reverse CCCGCTCGAG-TCCCAGCCCCGTGGAG	NdeI Yho!
		XhoI

272	Forward	AAAGAATTC-ATGACCGCAAAGGAAGAACTGTTCGC	
212			Eco RI
272		AAACTGCAG-TCAGAGCAGTTCCAAATCGGGGCT	Pst I
2/3		AAAGAATTC-ATGAGTCTTCAGGCGGTATTTATATACCC	Eco RI
274		AAACTGCAG-TTACGCGTAAGAAAAACTGC	Pst I
2/4	rorward	CGCGGATCCCATATG-ACAGATTTGGTTACGGAC	BamHI-
	Daviorea	CCCGCTCGAG-TTTGCTTTCAGTATTATTGAA	NdeI
276			XhoI
270	roiwatu	AAAAAAGAATTC- ATGATTTTGCCGTCCATCACGATGATGCG	Eco RI
	Reverse	AAAAAACTGCAG-CTACACCACCATCGGCGAATTTATGGC	<b>.</b>
277	Forward	AAAAAAGAATTC-ATGCCCCGCTTTGAGGACAAGCTCGTAGG	Pst I
211	Peverse	AAAAAACTGCAG-TCATAAGCCATGCTTACCTTCCAACAA	
2772	Forward	AAAAAAGAATTC-GGGGCGGCGGCTGGGTTGGACGTAGG	Pst I
211a	Peverce	AAAAAACTGCAG-TCATAAGCCATGCTTACCTTCCAACAA	Eco RI
278	Forward	AAAAAAGGTACC-GTCAAAGTTGTATTAATCGGGCCTTTGCC	Pst I
270	Reverse	AAAAAACTGCAG-TCATTCAACCATATCAAATCTGCC	Kpn I
2782		AAAAAAGAATTC-AAAACTCTCCTAATTCGTCATAGTCG	Pst I
2704	Deverce	AAAAAACTGCAG-TCATTCAACCATATCAAATCTGCC	Eco RI
279		CGCGGATCCCATATG-TTGCCTGCAATCACGATT	Pst I
219	rorward	COCOGNICCCATATO-TIGCCIGCAATCACGATI	BamHI-
	Reverse	CCCGCTCGAG-TTTAGAAGCGGGCGCAA	NdeI
280		AAAAAAGGTACC-GCCCCCTGCCGGTTGTAACCAG	XhoI
		AAAAAACTGCAG-TTATTGCTTCATCGCGTTGGTCAAGGC	Kpn I Pst I
281		AAAAAAGAATTC-GCACCCGTCGGCGTATTCCTCGTCATGCG	Eco RI
	Reverse	AAAAAATCTAGA-GGTCAGAATGCCGCCTTCTTTGCCGAG	Xba I
281a	Forward	AAAAAAGAATTC-TCCTACCACATCGAAATTCCTTCCGG	Eco RI
	Reverse	AAAAAATCTAGA-GGTCAGAATGCCGCCTTCTTTGCCGAG	Xba I
282		AAAAAAGAATTC-CTTTACCTTGACCTGACCAACGGGCACAG	Eco RI
	Reverse	AAAAAACTGCAG-TCAACCTGCCAGTTGCGGGAATATCGT	Pst I
283		CGCGGATCCCATATG-GCCGTCTTTACTTGGAAG	BamHI-
		occion mendamo	Ndel
	Reverse	CCCGCTCGAG-ACGGCAGTATTTGTTTACG	XhoI
284		CGCGGATCCCATATG-TTTGCCTGCAAAAGAATCG	BamHI-
			NdeI
		CCCGCTCGAG-CCGACTTTGCAAAAACTG	XhoI
286	Forward	CGCGGATCCCATATG-GCCGACCTTTCCGAAAA	BamHI-
	D	COCCCTCC A C C A A CCCCCTTC TO THE	NdeI
207		CCCGCTCGAG-GAAGCGCGTTCCCAAG	XhoI
287	rorward	CCGGAATTCTAGCTAGC-CTTTCAGCCTGCGGG	EcoRI-
	Reverse	CCCGCTCGAG-ATCCTGCTCTTTTTTGCC	NheI
288		CGCGGATCCCATATG-CACACCGGACAGG	XhoI
200	TOTWALL	COCOGATECCATATO-CACACCOGACAGO	BamHI-
	Reverse	CCCGCTCGAG-CGTATCAAAGACTTGCGT	Ndel
<b>2</b> 90		CGCGGATCCCATATG-GCGGTTTGGGGCGGA	XhoI
			BamHI-

294	Reverse Forward Reverse Forward Reverse	CCCGCTCGAG-TCGGCGCGGGGCCCCCCCCGCGGATCCCATATG-TGCGGGCAAACGCCCCCCCCGCTCGAG-TTGATTTTTGCGGATGATTTAAAAAAAAAA	NdeI XhoI BamHI- NdeI XhoI Eco RI Pst I
294	Reverse Forward Reverse Forward Reverse	CGCGGATCCCATATG-TGCGGGCAAACGCCC  CCCGCTCGAG-TTGATTTTTGCGGATGATTT  AAAAAAGAATTC-GTCTGGTCGATTCGGGTTGTCAGAAC  AAAAAACTGCAG-TTACCAGCTGATATAAAACATCGCTTT	BamHI- NdeI XhoI Eco RI
294	Reverse Forward Reverse Forward	CCCGCTCGAG-TTGATTTTTGCGGATGATTT  AAAAAAGAATTC-GTCTGGTCGATTCGGGTTGTCAGAAC  AAAAAACTGCAG-TTACCAGCTGATATAAAACATCGCTTT	NdeI XhoI Eco RI
295	Forward Reverse Forward	AAAAAAGAATTC-GTCTGGTCGATTCGGGTTGTCAGAAC AAAAAACTGCAG-TTACCAGCTGATATAAAACATCGCTTT	XhoI Eco RI
295	Forward Reverse Forward	AAAAAAGAATTC-GTCTGGTCGATTCGGGTTGTCAGAAC AAAAAACTGCAG-TTACCAGCTGATATAAAACATCGCTTT	Eco RI
295	Reverse Forward Reverse	AAAAAACTGCAG-TTACCAGCTGATATAAAACATCGCTTT	
	Forward Reverse	CGCGGATCCCATATG-AACCGGCCGGCCTCC	Pst I
	Reverse	COCOGATCCCATATO-AACCGGCCGGCCTCC	
	Reverse		BamHI-
	T 1	CCCGCTCGAG-CGATATTTGATTCCGTTGC	NdeI
297	rorward	AAAAAAGAATTC-GCATACATTGCTTCGACAGAGAG	XhoI
	Reverse	AAAAAACTGCAG-TCAATCCGATTGCGACACGGT	Eco RI
298	Forward	AAAAAAGAATTC-CTGATTGCCGTGTGGTTCAGCCAAAACCC	Pst I
	Reverse	AAAAAACTGCAG-TCATGGCTGTGTTGCTTGATGGTTGCGT	Eco RI
299	Forward	CGCGGATCCGCTAGC-CTACCTGTCGCCTCCG	Pst I
	1 O1 Wald	ede <u>ddateederade</u> -cracerdregeereeg	BamHI-
	Reverse	CCCGCTCGAG-TTGCCTGATTGCAGCGG	NheI
302	Forward	AAAAAAGAATTC-ATGAGTCAAACCGATACGCAACG	XhoI
	Reverse	AAAAAACTGCAG-TTAAGGTGCGGATAGAATGTGGGCGC	Eco RI
305	Forward	AAAAAAGGTACC-GAATTTTACCGATTTCCAGCACCGGA	Pst I
	Reverse	AAAAAACTGCAG-TCATTCCCAACTTATCCAGCCTGACAG	Kpn I
305a 1	Forward	AAAAAAGGTACC-TCCCGTTCGGGCAGTACGATTATGGG	Pst I
]	Reverse	AAAAAACTGCAG-TTACAAACCGACATCATGCAGGGTGAA	Kpn I
306 I	Forward	CGCGGATCCCATATG-TTTATGAACAAATTTTCCC	Pst I BamHI-
			Ndel
]	Reverse	CCCG <u>CTCGAG</u> -CCGCATCGGCAGAC	XhoI
308 I	Forward	CGCGGATCCCATATG-TTAAATCGGGTATTTTATC	BamHI-
			Ndel
]	Reverse	CCCGCTCGAG-ATCCGCCATTCCCTGC	XhoI
311 F	Forward .	AAAAAAGGTACC-ATGTTCAGTTTTGGCTGGGTGTTT	Kpn I
I	Reverse .	AAACTGCAG-ATGTTCATATTCCCTGCCTTCGGC	Pst I
312 F	Forward .	AAAAAGGTACC-ATGAGTATCCCATCCGGCGAAATT	Kpn I
ŀ	Reverse .	AAACTGCAG-TCAGTTTTTCATCGATTGAACCGG	Pst I
313 F	Forward .	AAAAAAGAATTC-ATGGACGACCCGCGCACCTACGGATC	Eco RI
F	Reverse .	AAAAAA <u>CTGCAG</u> -TCAGCGGCTGCCGCCGATTTTGCT	Pst I
401 F	Forward (	CGCGGATCCCATATG-AAGGCGGCAACACAGC	BamHI-
-	D	00000mgg	NdeI
402 T	keverse (	CCCGCTCGAG-CCTTACGTTTTTCAAAGCC	XhoI
402 F	Portugard A	AAAAAAGAATTC-GTGCCTCAGGCATTTTCATTTACCCTTGC	Eco RI
400 - T	keverse	AAAAAATCTAGA-TTAAATCCCTCTGCCGTATTTGTATTC	Xba I
402a F	Powerd A	AAAAAAGAATTC-AGGCTGATTGAAAACAAACACGG	Eco RI
406 E	Reverse A	AAAAAATCTAGA-TTAAATCCCTCTGCCGTATTTGTATTC	Xba I
406 F	orward (	CGCGGATCCCATATG-TGCGGGACACTGACAG	BamHI-
10	Reverse /	CCCCCTCGAC ACCTTCTCCTCCTTCTCTCTCTCTCTCTCTCTCTCTC	NdeI
	CVC13E (	CCCGCTCGAG-AGGTTGTCCTTGTCTATG	XhoI

501	Forward CGCGGATCCCATATG-GCAGGCGGAGATGGC	BamHI-
	Tormer of a desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired desired de	Ndel
	Reverse CCCGCTCGAG-GGTGTGATGTTCACCC	XhoI
502	Forward CGCGGATCCCATATG-GTAGACGCGCTTAAGCA	BamHI-
	omeneded hadea	Ndel
	Reverse CCCGCTCGAG-AGCTGCATGGCGGCG	
503		XhoI
303	Towns occommocontino-torreddddaaaddcg	BamHI-
	Reverse CCCGCTCGAG-CCGCGCATTCCTCGCA	NdeI
504	Forward CGCGGATCCCATATG-AGCGATATTGAAGTGACG	XhoI
504	Toward Cocoonfeccatato-Adedatatioaadidaed	BamHI-
	Reverse GCCCAAGCTT-TGATTCAAGTCCTTGCCG	NdeI
505	Forward CGCGGATCCCATATG-TTTCGTTTACAATTCAGG	HindIII
505	Forward COCOGATECCATATO-TITEGITTACAATTCAGG	BamHI-
	Reverse CCCGCTCGAG-CGGCGTTTTATAGCGG	NdeI
510	Forward CGCGGATCCCATATG-CCTTCGCGGACAC	XhoI
310	Forward COCOGATECCATATO-CCTTCGCGGACAC	BamHI-
	Reverse CCCGCTCGAG-GCGCACTGGCAGCG	Ndel
512	Forward CGCGGATCCCATATG-GGACATGAAGTAACGGT	XhoI
312	roiwaid CGCGGATCCCATATG-GGACATGAAGTAACGGT	BamHI-
	Reverse CCCGCTCGAG-AGGAATAGCCTTTGACG	Ndel
515		XhoI
313	Forward CGCGGATCCCATATG-GAGGAAATAGCCTTCGA	BamHI-
	Reverse CCCGCTCGAG-AAATGCCGCAAAGCATC	NdeI
516		XhoI
310	Forward CGCGGATCCCATATG-TGTACGTTGATGTTGTGG	BamHI-
	Reverse CCCGCTCGAG-TTTGCGGGCGGCATC	NdeI
517		XhoI
317	Forward CGCGGATCCCATATG-GGTAAAGGTGTGGAAATA	BamHI-
	Reverse CCCGCTCGAG-GTGCGCCCAGCCGT	NdeI
512	Forward AAAGAATTC-GCTTTTTTACTGCTCCGACCGGAAGG	XhoI
310		Eco RI
£10	Reverse AAACTGCAG-TCAAATTTCAGACTCTGCCAC	Pst I
319	Forward CGCGGATCCCATATG-TTCAAATCCTTTGTCGTCA	BamHI-
	Poverse CCCCCTCCAC TTTCCCCCCTTTTTCCCCC	NdeI
<b>530</b>	Reverse CCCGCTCGAG-TTTGGCGGTTTTGCTGC	XhoI
520	Forward CGCGGATCCCATATG-CCTGCGCTTCTTTCA	BamHI-
	Posterno CCCCCTCCAC ATATTTACATTTACA	NdeI
501	Reverse CCCGCTCGAG-ATATTTACATTTCAGTCGGC	XhoI
521	Forward CGCGGATCCCATATG-GCCAAAATCTATACCTGC	BamHI-
	Persona CCCCCTCGAG CATACCGGGGA COMPGG	NdeI
522	Reverse CCCGCTCGAG-CATACGCCCCAGTTCC	XhoI
522	Forward CGCGGATCCCATATG-ACTGAGCCGAAACAC	BamHI-
	Poverne CCCCA A CCTT TTCTC A TTTCTC	NdeI
522	Reverse GCCCAAGCTT-TTCTGATTTCAAATCGGCA	HindIII
523	Forward CGCGGATCCCATATG-GCTCTGCTTTCCGCG	BamHI-
		NdeI

	Reverse CCCGCTCGAG-AGGGTGTGTGATAATAAGAAG	
525		XhoI
1 323	Forward CGCGGATCCCATATG-GCCGAAATGGTTCAAATC	BamHI-
	Reverse CCCGCTCGAG-GCCCGTGCATATCATAAA	NdeI
527	Forward AAAGAATTC TTCCCTCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCTTCCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCCTTCAATCTTCCTTCCAATCTTCCTTCCAATCTTCCCTTCAATCTTCCTTCCAATCTTCCTTCCAATCTTCCTTCAATCTTCCTTCCAATCTTCCTTCAATCTTCCTTCAATCTTCCTTCCAATCTTCCTTCCAATCTTCCTTCAATCTTCCTTCAATCTTCCTTCAATCTTCCTTCAATCTTCCTTCAATCTTCCTTCAATCTTCCTTCAATCTTCCTTCAATCTTCCTTCAATCTTCCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCA	XhoI
1 327	Torward Thursday Tree Creating Toccording Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina Toccordina To	Eco RI
529	Reverse AAACTGCAG-TTATGCTAAACTCGAAACAAATTC	Pst I
329	Forward CGCGGATCCGCTAGC-TGCTCCGGCAGCAAAAC	BamHI-
	Reverse GCCCAAGCTT-ACGCAGTTCGGAATGGAG	NheI
530		HindIII
350	Forward CGCGGATCCCATATG-AGTGCGAGCGCGG	BamHI-
	Reverse CCCGCTCGAG-ACGACCGACTGATTCCG	NdeI
531		XhoI
	Reverse AAAAAACTGCAG-TTAAAACAGCGCCGTGCCGACGACAAG	Eco RI
532	Forward AAAAAAGAATTC-ATGAGCGGTCAGTTGGGCAAAGGTGC	Pst I
	Reverse AAAAAACTGCAG-TCAGTGTTCCAAGTGGTCGGTATCAAA	Eco RI
532a	Forward AAAAAAGAATTC-TTGGGTGTCGCGTTTGAGCCGGAAGT	Pst I
0020	Reverse AAAAAACTGCAG-TCAGTGTTCCAAGTGGTCGCTATCAAA	Eco RI
535	Forward AAAGAATTC-ATGCCCTTTCCCGTTTTCAGAC	Pst I
	Reverse AAACTGCAG-TCAGACGACCCCGCCTTCCCC	Eco RI
537	Forward CGCGGATCCCATATG-CATACCCAAAACCAATCC	Pst I
33,	Torward COCOGATCCCATATO-CATACCCAAAACCAATCC	BamHI-
	Reverse CCCGCTCGAG-ATCCTGCAAATAAAGGGTT	NdeI
538	Forward CGCGGATCCCATATG-GTCGAGCTGGTCAAAGC	XhoI
	dredAderdoreAAde	BamHI-
	Reverse CCCGCTCGAG-TGGCATTTCGGTTTCGTC	NdeI XhoI
539	Forward CGCGGATCCGCTAGC-GAGGATTTGCAGGAAA	Alloi BamHI-
		NheI
	Reverse CCCGCTCGAG-TACCAATGTCGGCAAATC	XhoI
542	Forward AAAGAATTC-ATGCCGTCTGAAACCGTGTC	Eco RI
	Reverse AAACTGCAG-TTACCGCGAACCGGTCAGGAT	Pst I
543	Forward AAAAAAAATTC-GCCTTCGATGGCGACGTTGTAGGTAC	Eco RI
	Reverse AAAAAATCTAGA-	Xba I
	TTAATGAAGAACATATTGGAATTTTGG	Abai
543a	Forward AAAAAAGAATTC-GGCAAAACTCGTCATGAATTTGC	Eco RI
	Reverse AAAAAATCTAGA-	Xba I
E 1 1	TTAATGAAGAAGAACATATTGGAATTTTGG	
3 <del>44</del>	Forward AAAGAATTC-GCGCCCGCCTTCTCCCTGCCCGACCTGCACGG	Eco RI
E A A =	Reverse AAACTGCAG-CTATTGCGCCACGCGCGTATCGAT	Pst I
3 <del>44</del> a	Forward AAAAAGAATTC-	Eco RI
	GCAAATGACTATAAAAACAAAAACTTCCAAGTACTTGC	
547	Reverse AAACTGCAG-CTATTGCGCCACGCGCGTATCGAT  Forward AAACAATTC ATCTTGCTA CATTAGAT	Pst I
J <del>-1</del> /	Forward AAAGAATTC-ATGTTCGTAGATAACGGATTTAATAAAAC	Eco RI
548	Reverse AAACTGCAG-TTAACAACAAAAAAAAAACGCTT	Pst I
J-0	Forward AAAGAATTC-GCCTGCAAACCTCAAGACAACAGTGCGGC	Eco RI

	Reverse	AAACTGCAG-TCAGAGCAGGGTCCTTACATCGGC	D-4 T
550		AAAAAGTCGAC-	Pst I
	101	ATGATAACGGACAGGTTTCATCTCTTTCATTTTCC	Sal I
	Reverse	AAACTGCAG-TTACGCAAACGCTGCAAAATCCCC	Pst I
550a		AAAAAGAATTC-GTAAATCACGCCTTTGGAGTCGCAAACGG	Eco RI
	Reverse	AAACTGCAG-TTACGCAAACGCTGCAAAATCCCC	Pst I
552		AAAAAGAATTC-TTGGCGCGTTGGCTGGATAC	
332		AAACTGCAG-TTATTTCTGATGCCTTTTCCCAAC	Eco RI
554		CGCGGATCCCATATG-TCGCCCGCGCCCAAC	Pst I
334	TOIWAIG	ede <u>ddateeeatatd</u> -tedeeededeecaac	BamHI-
	Reverse	CCCGCTCGAG-CTGCCCTGTCAGACAC	NdeI
556		AAAGAATTC-GCGGGCGGTTTTGTTTGGACATCCCG	XhoI
		AAACTGCAG-TTAACGGTGCGGACGTTTCTGACC	Eco RI
557		CGCGGATCCCATATG-TGCGGTTTCCACCTGAA	Pst I
337	1 of ward	ede <u>ddateeeatatu</u> -tdeddiffeeaccigaa	BamHI-
	Reverse	CCCGCTCGAG-TTCCGCCTTCAGAAAGG	NdeI
558		AAAGAATTC-GAGCTTTATATGTTTCAACAGGGGACGGC	XhoI
		AAACTGCAG-CTAAACAATGCCGTCTGAAAGTGGAGA	Eco RI
558a		AAAAAAGAATTC-ATTAGATTCTATCGCCATAAACAGACGGG	Pst I
		AAAAAACTGCAG-CTAAACAATGCCGTCTGAAAGTGGAGA	
560		AAAAAGAATTC-	Pst I
	101111111	TCGCCTTTCCGGGACGGGGCGCACAAGATGGC	Eco RI
	Reverse	AAAAAACTGCAG-TCATGCGGTTTCAGACGGCATTTTGGC	Pst I
561	Forward	CCGGAATTCTACATATG-ATACTGCCAGCCCGT	EcoRI-
			NdeI
	Reverse	CCCGCTCGAG-TTTCAAGCTTTCTTCAGATG	XhoI
562	Forward	CGCGGATCCCATATG-GCAAGCCCGTCGAG	BamHI-
			NdeI
		CCCG <u>CTCGAG</u> -AGACCAACTCCAACTCGT	XhoI
565	Forward	CGCGGATCCCATATG-AAGTCGAGCGCGAAATAC	BamHI-
			NdeI
		CCCGCTCGAG-GGCATTGATCGGCGGC	XhoI
566	Forward	CGC <u>GGATCCCATATG</u> -GTCGGTGGCGAAGAGG	BamHI-
	Darrana	CCCCCTCC A C CCCATTCCCCCC + + === :	NdeI
E 67		CCCGCTCGAG-CGCATGGGCGAAGTCA	XhoI
567	rorward	CCGGAATTCTACATATG-AGTGCGAACATCCTTG	EcoRI-
	Reverse	CCCGCTCGAG-TTTCCCCGACACCCTCG	NdeI
568		CGCGGATCCCATATG-CTCAGGGTCAGACC	XhoI
200	TOIWAIU	COCOGATECCATATO-CTCAGGGTCAGACC	BamHI-
	Reverse	CCCGCTCGAG-CGGCGCGCGTTCAG	NdeI
		AAAAAGAATTC-CTGATTGCCTTGTGGGAATATGCCCG	XhoI
	Reverse	AAAAAACTGCAG-TTATGCATAGACGCTGATAACGGCAAT	Eco RI
570	Forward	CGCGGATCCCATATG-GACACCTTCCAAAAAATCG	Pst I
2.0			BamHI-
	Reverse	CCCGCTCGAG-GCGGCGTTCATTTCTTT	NdeI XhoI
			VIIOI

571	Forward AAAAAAGAATTC-	
	ATGGGTATTGCCGGCGCCGTAAATGTTTTGAACCC	Eco RI
	Reverse AAAAAACTGCAG-TTATGGCCGACGCGCGCTACCTGACG	<b>.</b>
572	Forward CGCGGATCCCATATG-GCGCAAAAAGGCAAAACC	Pst I
	TOWNED GOOGLATIO-OCOCAAAAGGCAAAACC	BamHI-
	Reverse CCCGCTCGAG-GCGCAGTGTGCCGATA	NdeI
573		XhoI
	- The design of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of the second of	BamHI-
	Reverse CCCGCTCGAG-GACGGTGTCATTTCGCC	NdeI
574		XhoI
		BamHI-
	Reverse CCCGCTCGAG-AACTTCGATTTTATTCGGG	NdeI
575		XhoI
		BamHI-
	Reverse CCCGCTCGAG-CATTCCGAATCTGAACAG	NdeI XhoI
576		Alloi BamHI-
		NdeI
ĺ	Reverse CCCGCTCGAG-ATTTACTTTTTTGATGTCGAC	XhoI
577	Forward CGCGGATCCCATATG-GAAAGGAACGGTGTATTT	BamHI-
		NdeI
	Reverse CCCGCTCGAG-AGGCTGTTTGGTAGATTCG	XhoI
578	Forward CGCGGATCCCATATG-AGAAGGTTCGTACAG	BamHI-
		NdeI
	Reverse CCCGCTCGAG-GCCAACGCCTCCACG	XhoI
579	Forward CGCGGATCCCATATG-AGATTGGGCGTTTCCAC	BamHI-
	Payarra CCCCCTCCAC ACAATTTCATTCATTCATTCATTCATTC	NdeI
580	Reverse CCCGCTCGAG-AGAATTGATGATGTGTATGT	XhoI
360	Forward CGCGGATCCCATATG-AGGCAGACTTCGCCGA	BamHI-
	Reverse CCCGCTCGAG-CACTTCCCCCGAAGTG	NdeI
581	Forward CGCGGATCCCATATG-CACTTCGCCCAGC	XhoI
	Toward Cocoda recentato-caci regeccage	BamHI-
	Reverse CCCGCTCGAG-CGCCGTTTGGCTTTGG	NdeI
582	Forward AAAAAAGAATTC-TTTGGAGAGACCGCGCTGCAATGCGC	XhoI
	Reverse AAAAAATCTAGA-TCAGATGCCGTCCCAGTCGTTGAA	Eco RI
583	Forward AAAAAAGAATTC-ACTGCCGGCAATCGACTGCATAATCG	Xba I
	Reverse AAAAAACTGCAG-TTAACGGAGGTCAATATGATGAAATTG	Eco RI
584	Forward AAAAAAGAATTC-	Pst I
	GCGGCTGAAGCATTGAATTACAATATTGTC	Eco RI
	Reverse AAAAAACTGCAG-TCAGAACTGAACCGTCCCATTGACGCT	D-4 T
585	Forward AAAAAAGGTACC-TCTTTCTGGCTGCTGCAGAACACCCTTGC	Pst I
	Reverse AAAAAACTGCAG-TCAGTTCGCACTTTTTCTGTTTTGGA	Eco RI
586	Forward CGCGGATCCCATATG-GCAGCCCATCTCG	Pst I
		BamHI-
	Reverse CCCGCTCGAG-TTTCAGCGAATCAAGTTTC	NdeI
587	Forward CGCGGATCCCATATG-GACCTGCCCTTGACGA	XhoI
		BamHI-

j	Darrage	CCCCTCCAC AAATCTATCCTCTACCCC	NdeI
500		CCCGCTCGAG-AAATGTATGCTGTACGCC	XhoI
588		AAAAAAGAATTC-GCCGTCCTGACTTCCTATCAAGAACCAGG	Eco RI
		AAAAAACTGCAG-TTATTTGTTTTTTGGGCAGTTTCACTTC	Pst I
589	Forward	AAAAAAGAATTC-	Eco RI
	<b>D</b>	ATGCAACAAAAATCCGTTTCCAAATCGAAGG	
500		AAAAAACTGCAG-CTAATCGATTTTTACCCGTTTCAGGCG	Pst I
590		AAAAAAGAATTC-ATGAAAAAACCTTTGATTTCAGTTGCGGC	Eco RI
		AAAAAACTGCAG-TTACTGCTGCGGCTCTGAAACCAT	Pst I
591		AAAAAAGAATTC-CACTACATCGTTGCCAGATTGTGCGG	Eco RI
		AAAAAACTGCAG-CTAACCGAGCAGCCGGGTAACGTCGTT	Pst I
592a	Forward	AAAAAAGAATTC-CGCGATTACACCGCCAAGCTGAAAATGGG	Eco RI
		AAAAAACTGCAG-TTACCAAACGTCGGATTTGATACG	Pst I
593	Forward	CGCGGATCCGCTAGC-CTTGAACTGAACGGACTC	BamHI-
	D	CCCCCTCC + C CCCC + + CCCC + CC	NheI
504		CCCGCTCGAG-GCGGAAGCGGACGATT	XhoI
594a		AAAAAAGAATTC-GGTAAGTTCGCCGTTCAGGCCTTTCA	Eco RI
		AAAAAACTGCAG-TTACGCCGCCGTTTCCTGACACTCGCG	Pst I
595		AAAAAAGAATTC-TGCCAGCCGCCGGAGGCGGAGAAAGC	Eco RI
		AAAAAACTGCAG-TTATTTCAAGCCGAGTATGCCGCG	Pst I
596	Forward	CGCGGATCCCATATG-TCCCAACAATACGTC	BamHI-
	<b>D</b>	COCCOTOCA C. A COCCOTTA COCCOTTA	NdeI
507		CCCGCTCGAG-ACGCGTTACCGGTTTGT	XhoI
397	rorward	CGCGGATCCCATATG-CTGCTTCATGTCAGC	BamHI-
	Reverse	GCCCAAGCTT-ACGTATCCAGCTCGAAG	Ndel
601		CGCGGATCCCATATG-ATATGTTCCCAACCGGCAAT	HindIII
001	1 OI Walu	CGCGGGATCCCATATG-ATATGTTCCCAACCGGCAAT	BamHI-
	Reverse	CCCGCTCGAG-AAAACAATCCTCAGGCAC	NdeI XhoI
602		CGCGGATCCGCTAGC-TTGCTCCATCAATGC	BamHI-
		ridereeminge	NheI
	Reverse	CCCGCTCGAG-ATGCAGCTGCTAAAAGCG	XhoI
603	Forward	AAAAAAGAATTC-CTGTCCTCGCGTAGGCGGGGACGGG	Eco RI
		AAAAAACTGCAG-CTACAAGATGCCGGCAAGTTCGGC	Pst I
604		CGCGGATCCGCTAGC-CCCGAAGCGCACTT	BamHI-
			NheI
	Reverse	CCCG <u>CTCGAG</u> -GACGGCATCTGCACGG	XhoI
606a	Forward	AAAAAAGAATTC-CGCGAATACCGCGCCGATGCGGGCGC	Eco RI
	Reverse	A A A A A A COTOCO A COMPANA A COCO A COTOCO A C	Pst I
607		A A A A A C A A TTTC A TCCTCCTCCTCCC A CCTC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGACGGCCTTATGCGATCTGAC	Pst I
608		A A A A A C A A TTC A TCTCCCCCCCCCCCCCC	Eco RI
	Reverse	A A A A A CTCC A C TT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A CTCT A	Pst I
609		CCCCC ATCCCCATATC CTTCTCCCATACACTCC	BamHI-
			Ndel

	D	CCCCCTCC A C CTCC A TO CTCC	
610	Keverse	CCCGCTCGAG-CTGGATTATGATGTCTGTC	XhoI
010	rorward	CGCGGATCCCATATG-ATTGGAGGGCTTATGCA	BamHI-
	Dorross	CCCCCTCC AC ACCCTTCA ACCTTCA	NdeI
611		CCCGCTCGAG-ACGCTTCAACATCTTTGCC	XhoI
611	Forward	CGCGGATCCCATATG-CCGTCTCAAAACGGG	BamHI-
	Dayaraa	CCCCCTCCAC AAGGAGTTTTGAAAGGAGTTTTGAAAGGAGTTTTGAAAGGAGTTTTTGAAAGGAGTTTTTGAAAGGAGTTTTTGAAAGGAGTTTTTGAAAGGAGTTTTTT	NdeI
613	Reverse	CCCGCTCGAG-AACGACTTTGAACGCGCAA	XhoI
013	rorward	CGCGGATCCCATATG-TCGCGTTCGAGCCG3	BamHI-
	Deverse	CCCCCTCC A C A CCCTCT A A A TO A TO A TO	NdeI
614	Economic	CCCGCATCCCATATC TOCCTOR	XhoI
014	roiwaiu	CGC <u>GGATCCCATATG</u> -TCCGTCGTGAGCGGC	BamHI-
	Reverse	CCCGCTCGAG-CCATACTGCGGCGTTC	NdeI
616	Forward	AAAAAAAATTC ATCTCAAACAAATTC	XhoI
010	Deverse	AAAAAAGAATTC-ATGTCAAACACAATCAAAATGGTTGTCGG	Eco RI
6100	Keverse	AAAAAATCTAGA-TTAGTCCGGGCGGCAGGCAGCTCG	Xba I
UIYA	Powerd .	AAAAAAGAATTC-GGGCTTCTCGCCGCCTCGCTTGC	Eco RI
622	Reverse .	AAAAAACTGCAG-TCATTTTTTGTGTTTTAAAACGAGATA	Pst I
022	Forward	CGCGGATCCCATATG-GCCGCCCTGCCTAAAG	BamHI-
	Deverse		NdeI
624	Formerd	CCCGCATCCCATATC TOCCCAAATGATAAATCTG	XhoI
024	rorwaru	CGCGGATCCCATATG-TCCCCGCGCTTTTACCG	BamHI-
	Reverse	CCCG <u>CTCGAG</u> -AGATTCGGGCCTGCGC	Ndel
625			XhoI
023	Torward	CGC <u>GGATCCCATATG</u> -TTTGCAACCAGGAAAATG	BamHI-
	Reverse (	CCCGCTCGAG-CGGCAAAATTACCGCCTT	NdeI
627a		AAAAAAGAATTC-AAAGCAGGCGAGGCAGGCGCGCTGGG	XhoI
	Reverse	AAAAACTGCAG-	Eco RI
		TTACGAATGAAACAGGGTACCCGTCATCAAGGC	Pst I
628	Forward A	AAAAAAGGTACC-GCCTTACAAACATGGATTTTGCGTTC	Tr - T
	Reverse A	AAAAAACTGCAG-CTACGCACCTGAAGCGCTGGCAAA	Kpn I
629a	Forward A	AAAAAAGAATTC-GCCACCTTTATCGCGTATGAAAACGA	Pst I
	Reverse /	AAAAAACTGCAG-TTACAACACCGCCGTCCGGTTCAAACC	Eco RI
630a	Forward A	AAAAAAGAATTC-GCGGCTTTGGGTATTTCTTTCGG	Pst I
0004	Reverse A	AAAAAACTGCAG-TTAGGAGACTTCGCCAATGGAGCCGGG	Eco RI
635	Forward A	AAAAAGAATTC-	Pst I
000		ATGACCCAGCGACGGGTCGGCAAGCAAAACCG	Eco RI
	Reverse A	AAAAACTGCAG-TTAATCCACTATAATCCTGTTGCT	<b>.</b>
638	Forward A	AAAAAGAATTC-ATGATTGGCGAAAAGTTTATCGTAGTTGG	Pst I
	Reverse A	AAAAACTGCAG-TCACGAACCGATTATGCTGATCGG	Eco RI
639	Forward (	CGCGGATCCCATATG-ATGCTTATTTTGTTCG	Pst I
	_ O. Watu C	CONTROCATATO-AIGCITTATTTTGTTCG	BamHI-
	Reverse (	CCCGCTCGAG-ATCGCGGCTGCCGAC	Ndel
642	Forward C	CGCGGATCCCATATG-CGGTATCCGCCGCAAT	XhoI
		COSTITUCCATATO-COGTATCCGCGCAAT	BamHI-
	Reverse C	CCGCTCGAG-AGGATTGCGGGGCATTA	NdeI
			XhoI

643	Forward	CGCGGATCCCATATG-GCTTCGCCGTCGGCAG	BamHI-
		de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constante de la constant	NdeI
	Reverse	CCCGCTCGAG-AACCGAAAAACAGACCGC	XhoI '
644		AAAAAAGAATTC-	Eco RI
		ATGCCGTCTGAAAGGTCGGCGGATTGTTGCCC	LCO KI
	Reverse	AAAAAATCTAGA-CTACCCGCAATATCGGCAGTCCAATAT	Pst I
645		AAAAAGAATTC-GTGGAACAGAGCAACACGTTAAATCG	Eco RI
		AAAAACTGCAG-CTACGAGGAAACCGAAGACCAGGCCGC	Pst I
647		AAAAAGAATTC-ATGCAAAGGCTCGCCGCAGACGG	Eco RI
		AAAAACTGCAG-TTAGATTATCAGGGATATCCGGTAGAA	Pst I
648		AAAAAGAATTC-	Eco RI
		ATGAACAGGCGCGACGCGCGGATCGAACG	LCO KI
	Reverse	AAAAAACTGCAG-TCAAGCTGTGTGCTGATTGAATGCGAC	Pst I
649		AAAAAAGAATTC-GGTACGTCAGAACCCGCCCACCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAACGGCGGAAACTGCCGCCGTC	Pst I
650		AAAAAAGAATTC-ATGTCCAAAACTCAAAACCATCGC	Eco RI
		AAAAAACTGCAG-TCAGACGGCATGGCGGTCTGTTTT	Pst I
652		AAAAAGGTACC-	Kpn I
		GCTGCCGAAGACTCAGGCCTGCCGCTTTACCG	p
		AAAAAACTGCAG-TTATTTGCCCAGTTGGTAGAATGCGGC	Pst I
653		AAAAAAGAATTC-GCGGCTTTGCCGGTAATTTTCATCGG	Eco RI
•		AAAAAACTGCAG-CTATGCCGGTCTGGTTGCCGGCGGCGA	Pst I
656a		AAAAAAGAATTC-CGGCCGACGTCGTTGCGTCCTAAGTC	Eco RI
		AAAAAACTGCAG-CTACGATTTCGGCGATTTCCACATCGT	Pst I
657		AAAAAAGAATTC-GCAGAATTTGCCGACCGCCATTTGTGCGC	Eco RI
		AAAAAACTGCAG-TTATAGGGACTGATGCAGTTTTTTTGC	Pst $\mathbf{I}_{\sim i}$
658	Forward	CGCGGATCCCATATG-GTGTCCGGAATTGTG	BamĤI-
	<b>D</b>	CCCCCTCC A C CCC A C A P C T T T T T T T T T T T T T T T T T T	NdeI
		CCCGCTCGAG-GGCAGAATGTTTACCGTT	XhoI
661	Forward	AAAAAAGAATTC-	Eco RI
	Dovorco	ATGCACATCGGCGGCTATTTTATCGACAACCC	
663		AAAAAACTGCAGATATCTCTCTCTCTCTCGCCGTCGGGC	Pst I
003	1.01 Mard	CGCGGATCCCATATG-TGTATCGAGATGAAATT	BamHI-
	Reverse	CCCGCTCGAG-GTAAAAATCGGGGCTGC	Ndel
664		CGCGGATCCCATATG-GCGCTGGCGCGGT	XhoI
		<u></u> 000001000001	BamHI- NdeI
	Reverse	CCCGCTCGAG-AAATCGAGTTTTACACCAC	XhoI
665		AAAAAAGAATTC-ATGAAATGGGACGAAACGCGCTTCGG	Eco RI
		AAAAAACTGCAG-TCAATCCAAAATTTTGCCGACGATTTC	Pst I
666		AAAAAAGAATTC-AACTCAGGCGAAGGAGTGCTTGTGGC	Eco RI
	Reverse	AAAAAATCTAGA-TCAGTTTAGGGATAGCAGGCGTAC	Xba I
667		AAAAAAGAATTC-	Eco RI
		CCGCATCCGTTTGATTTCCATTTCGTATTCGTCCG	200 10
	Reverse	AAAAAACTGCAG-TTAATGACACAATAGGCGCAAGTC	Pst I

669	Forward AAAAAAGAATTC-ATGCGCCGCATCATTAAAAAACACCAGC	~ =
	Reverse AAAAAACTGCAG-TTACAGTATCCGTTTGATGTCGGC	
670	Forward AAAAAAGAATTC-AAAAACGCTTCGGGCGTTTCGTCTTC	Pst I
	Reverse AAAAAACTGCAG-	Eco RI
	TTAGGAGCTTTTGGAACGCGTCGGACTGGC	Pst I
671	Forward CGCGGATCCCATATG-ACCAGCAGGGTAAC	<b>T</b>
		BamHI-
	Reverse CCCGCTCGAG-AGCAACTATAAAAACGCAAG	NdeI XhoI
672		BamHI-
		NdeI
	Reverse CCCGCTCGAG-ACGGGATAGGCGGTTG	XhoI
673	Forward AAAAAAGAATTC-ATGGATATTGAAACCTTCCTTGCAGG	Eco RI
	Reverse AAAAAACTGCAG-CTACAAACCCAGCTCGCGCAGGAA	Pst I
674	Forward AAAAAAGAATTC-ATGAAAACAGCCCGCCGCCGTTCCCG	Eco RI
	Reverse AAAAAACTGCAG-TCAACGGCGTTTGGGCTCGTCGGG	Pst I
675	Forward CGCGGATCCCATATG-AACACCATCGCCCC	BamHI-
	D. GGGGGTGG A G THE COLOR	NdeI
C77	Reverse CCCGCTCGAG-TTCTTCGTCTTCAAACTGT	XhoI
6//a	Forward AAAAAAGAATTC-AGACGGCATTCCCGATCAGTCGATTTTGA	Eco RI
600	Reverse AAAAAACTGCAG-TTACGTATGCGCGAAATCGACCGCCGC	Pst I
080	Forward CGCGGATCCGCTAGC-ACGAAGGGCAGTTCGG	BamHI-
	Reverse CCCGCTCGAG-CATCAAAAACCTGCCGC	NheI
681		XhoI
001	Forward AAAAAAGAATTC-ATGACGACGCCGATGGCAATCAGTGC Reverse AAAAAACTGCAG-TTACCGTCTTCCGCAAAAAACAGC	Eco RI
683	Forward CGCGGATCCCATATG-TGCAGCACACCGGACAA	Pst I
005	TOWARD OGOGOTICCCATATO-TOCAGCACACCGGACAA	BamHI-
	Reverse CCCGCTCGAG-GAGTTTTTTCCGCATACG	NdeI
684	Forward CGCGGATCCCATATG-TGCGGTACTGTGCAAAG	XhoI
		BamHI- NdeI
	Reverse CCCGCTCGAG-CTCGACCATCTGTTGCG	Xhol
685	Forward CGCGGATCCCATATG-TGTTTGCTTAATAATAAACATT	BamHI-
		NdeI
	Reverse CCCGCTCGAG-CTTTTTCCCCGCCGCA	XhoI
<b>68</b> 6	Forward CGCGGATCCCATATG-TGCGGCGGTTCGGAAG	BamHI-
	Payrama CCCCCTCCAC CATTTCCCAA TOTACA	NdeI
687	Reverse CCCGCTCGAG-CATTCCGATTCTGATGAAG	XhoI
067	Forward CGCGGATCCCATATG-TGCGACAGCAAAGTCCA	BamHI-
	Reverse CCCGCTCGAG-CTGCGCGGCTTTTTGTT	Ndel
<b>69</b> 0	Forward CGCGGATCCCATATG-TGTTCTCCGAGCAAAGAC	XhoI
<del>-</del>	- TO SOM TO TO TO TO TO TO TO TO TO TO TO TO TO	BamHI-
	Reverse CCCGCTCGAG-TATTCGCCCCGTGTTTGG	Ndel
691	Forward CGCGGATCCCATATG-GCCACGGCTTATATCCC	XhoI
		BamHI- NdeI
	Reverse CCCGCTCGAG-TTTGAGGCAGGAAGAAG	XhoI
		71101

694	Forward	CGCGGATCCCATATG-TTGGTTTCCGCATCCGG	70 777
0,74	TOTWARA	- COCCATCCOCTITATO-TIGGTTTCCGCATCCGG	BamHI-
1	Reverse	CCCGCTCGAG-TCTGCGTCGGTGCGGT	Ndel
695		CGCGGATCCCATATG-TTGCCTCAAACTCGTCCG	XhoI
	1 Ol Walu	CGCGGATCCCATATG-TTGCCTCAAACTCGTCCG	BamHI-
	Reverse	CCCGCTCGAG-TCGTTTGCGCACGGCT	NdeI
606		CGCGGATCCCATATG-TTGGGTTGCCGGCAGG	XhoI
090	TOIWAIU	COCOGNICCCATATO-TIGGGTTGCCGGCAGG	BamHI-
	Reverse	CCCGCTCGAG-TTGATTGCCGCAATGATG	Ndel
7000			XhoI
/00a		AAAAAGAATTC-GCATCGACAGACGGTGTGTCGTGGAC	Eco RI
701		AAAAAACTGCAG-TTACGCTACCGGCACGACTTCCAAACC	Pst I
/01	Forward	CGCGGATCCCATATG-AAGACTTGTTTGGATACTTC	BamHI-
	D	CCCCCTCC AC TCCCC ACA ACA ACA ACA	NdeI
700		CCCGCTCGAG-TGCCGACAACAGCCTC	XhoI
702		AAAAAGAATTC-ATGCCGTGTTCCAAAGCCAGTTGGATTTC	Eco RI
		AAAAAACTGCAG-TTAACCCCATTCCACCCGGAGAACCGA	Pst I
703	Forward	CGCGGATCCGCTAGC-CAAACGCTGGCAACCG	BamHI-
	_		NheI
		CCCGCTCGAG-TTTTGCAGGTTTGATGTTTG	XhoI
704a		AAAAAAGAATTC-GCTTCTACCGGTACGCTGGCGCG	Eco RI
ĺ	Reverse	AAAAACTGCAG-	Pst I
		TTAGTTTTGCCGGATAATATGGCGGGTGCG	
707	Forward	CGCGGATCCGCTAGC-GAAATTATTAACGATGCAGA	BamHI-
	-	COCCOTTO A C C A A A TOTAL A C C A A A TOTAL A C C A A A A TOTAL A C C A A A A TOTAL A C A A A A A A A A A A A A A A A A A	NheI
700		CCCGCTCGAG-GAAACTGTAATTCAAGTTGA	XhoI
708	Forward	CGCGGATCCGCTAGC-CCTTTTAAGCCATCCAAAA	BamHI-
	D	CCCCCTCC A C TTC A CCCCTC A CCC	Nhel
710		CCCGCTCGAG-TTGACCGGTGAGGACG	Xhol
/10	Forward	CGCGGATCCCATATG-GAAACCCACGAAAAAATC	BamHI-
	Dawana	CCCCCTCC A C A A CCCTTTTTCCCTC A C	NdeI
714		CCCGCTCGAG-AACGGTTTCGGTCAG	XhoI
/14	rorward	CGCGGATCCCATATG-AGCTATCAAGACATCTT	BamHI-
	Dayoras	CCCCCTCCAC CCCCTACCTAAAATCCC	NdeI
716		CCCGCTCGAG-GCGGTAGGTAAATCGGAT	XhoI
/16	rorward	CGCGGATCCCATATG-GCCAACAACCGGCAAG	BamHI-
	Dayoraa	CCCCCTCC & C TTT & C & A CCCC A TTTT & C &	Ndel
710		CCCGCTCGAG-TTTAGAACCGCATTTGCC	XhoI
718	rorward	CGCGGATCCCATATG-GAGCCGATAATGGCAAA	BamHI-
	Davorco	CCCCCTCCAC CCCCCCCCCA TCCTCTTCTCC	NdeI
720		CCCGCTCGAG-GGCGCGGGCATGGTCTTGTCC	XhoI
120	rotward	CGCGGATCCCATATG-AGCGGATGGCATACC	BamHI-
	Reverse	CCCGCTCGAG-TTTTGCATAGCTGTTGACCA	NdeI
723			XhoI
123	rorward	CGCGGATCCCATATG-CGACCCAAGCCCC	BamHI-
	Reverse	CCCGCTCGAG-AATGCGAATCCGCCGCC	NdeI
L	1010100		XhoI

725	Forward CGCGGATCCCATATG-GTGCGCACGGTTAAA	
/23	Torward CGCGGATCCCATATG-GTGCGCACGGTTAAA	BamHI-
1	Reverse CCCGCTCGAG-TTGCTTATCCTTAAGGGTTA	NdeI
726	Forward CCCCCATCCCATATC ACCATCATCATC	XhoI
1/20	Forward CGCGGATCCCATATG-ACCATCTATTTCAAAAAC	BamHI-
	Persona CCCCCTCCAC CCCCATCTTTA CCCTT	NdeI
728	Reverse CCCGCTCGAG-GCCGATGTTTAGCGTCC	XhoI
120	Forward CGCGGATCCCATATG-TTTTGGCTGGGAACGGG	BamHI-
	Paraman COCCOTTOCA C CTCA CA A A COTTO	NdeI
720	Reverse CCCGCTCGAG-GTGAGAAAGGTCGCGC	XhoI
729	Forward CGCGGATCCCATATG-TGCACCATGATTCCCCA	BamHI-
	Daviera COCOA A COTTO TITO CITA CONTO	NdeI
721	Reverse GCCCAAGCTT-TTTGTCGGTTTTGGGTATC	HindIII
731	Forward CGCGGATCCGCTAGC-GCCGTGCCGGAGG	BamHI-
	D	NheI
722	Reverse CCCGCTCGAG-ACGGGCGCGCAG	XhoI
/32	Forward CCGGAATTCTACATATG-TCGAAACCTGTTTTTAAGAA	EcoRI-
	D	NdeI
722	Reverse CCCGCTCGAG-CTTCTTATCTTTTTATCTTTC	XhoI
733	Forward CGCGGATCCCATATG-GCCTGCGGCGGCAA	BamHI-
	D COCCOMOCA C MCCCOMO	NdeI
724	Reverse CCCGCTCGAG-TCGCTTGCCTCCTTTAC	XhoI
/34	Forward CGCGGATCCCATATG-GCCGATACTTACGGCTAT	BamHI-
	Description of COCCOTTOCA CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTROL OF A TOTAL CONTRO	NdeI
725	Reverse CCCGCTCGAG-TTTGAGATTTTGAATCAAAGAG	Xho $I$
133	Forward CGCGGATCCCATATG-AAGCAGCAGCGGTCA	BamHI-
	Payrama CCCCCTCCAC ATTTCCCCT ACCC	NdeI
737	Reverse CCCGCTCGAG-ATTTCCGTAGCCGAGGG	XhoI
/3/	Forward CGCGGATCCCATATG-CACCACGACGGACACG	BamHI-
	Payama CCCCCTCCAC CTCCTCCCCCCCC	NdeI
739	Reverse CCCGCTCGAG-GTCGTCGCGGGGA	XhoI
139	Forward CGCGGATCCCATATG-GCAAAAAAACCGAACA	BamHI-
	Perverse CCCCCTCCAC CAACACTTTTTTTTTTTTTTTTTTTTT	NdeI
740	Reverse CCCGCTCGAG-GAAGAGTTTGTCGAGAATT	XhoI
740	Forward CGCGGATCCCATATG-GCCAATCCGCCCGAAG	BamHI-
	Perverse CCCCCTCCAC AAACCCCCCAAAACCCCCCAAAACCCCCCCAAAACCCCCC	NdeI
741	Reverse CCCGCTCGAG-AAACGCGCCAAAATAGTG	XhoI
/41	Forward CGCGGATCCCATATG-TGCAGCAGCGGAGGG	BamHI-
	Persona CCCCCTCCAC TTCCTTCCCCCC	NdeI
743	Reverse CCCGCTCGAG-TTGCTTGGCGGCAAGGC	XhoI
743	Forward CGCGGATCCCATATG-GACGGTGTTGTGCCTGTT	BamHI-
	Payerra CCCCCTCCAC CTTACCCATTCAA	NdeI
745	Reverse CCCGCTCGAG-CTTACGGATCAAATTGACG	XhoI
143	Forward CGCGGATCCCATATG-TTTTGGCAACTGACCG	BamHI-
	Reverse CCCGCTCCAC CAAATCACATCAC	NdeI
746	Reverse CCCGCTCGAG-CAAATCAGATGCCTTTAGG	XhoI
/ <del>1</del> U	Forward CGCGGATCCCATATG-TCCGAAAACAAACAAAAC	BamHI-
		· —-—

	Doverse	CCCGCTCGAG-TTCATTCGTTACCTGACC	Ndel
747			XhoI
/4/	Forward	CCGGAATTCTAGCTAGC-CTGACCCCTTGGG	EcoRI-
	Darraga	CCCCA A CCTT TTTTC A TTTTA A TTTC A CT A TT C A CT A TT C A CT A TT C A CT A TT C A CT A TT C A CT A TT C A CT A TT C A CT A TT C A CT A TT C A CT A TT C A CT A TT C A CT A TT C A CT A TT C A CT A TT C A CT A TT C A CT A TT C A CT A TT C A CT A TT C A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A CT A	NheI
740		GCCCAAGCTT-TTTTGATTTTAATTGACTATAGAAC	HindIII
749	Forward	CGCGGATCCCATATG-TGCCAGCCGCCG	BamHI-
	Darrama	CCCCCTCC A C TTTC A A CCCC A CT A TCC	NdeI
750		CCCGCTCGAG-TTTCAAGCCGAGTATGC	XhoI
/30	Forward	CGCGGATCCCATATG-TGTTCGCCCGAACCTG	BamHI-
	Darrama	CCCCTCC & C CTTTTTCCCCCCCC	NdeI
750		CCCGCTCGAG-CTTTTTCCCCGCCGCAA	XhoI
/58	Forward	CGCGGATCCCATATG-AACAATCTGACCGTGTT	BamHI-
	Darramaa	CCCCCTCC A C TCCCTC A A TCCTTTTCTCC	NdeI
750		CCCGCTCGAG-TGGCTCAATCCTTTCTGC	XhoI
/39	rorward	CGCGGATCCGCTAGC-CGCTTCACACACACCAC	BamHI-
	Darrage		NheI
762		CCCGCTCGAG-CCAGTTGTAGCCTATTTTG	XhoI
/03	rorward	CGCGGATCCCATATG-CTGCCTGAAGCATGGCG	BamHI-
	Dovorco		NdeI
761		CCCGCATCCCATATC TTTTTTCCCCCAAATACCGTTTCC	XhoI
704	rorward	CGCGGATCCCATATG-TTTTTCTCCGCCCTGA	BamHI-
	Deverse	CCCGCTCGAG-TCGCTCCCTAAAGCTTTC	Ndel
765			XhoI
703	roiwaiu	CGCGGATCCCATATG-TTAAGATGCCGTCCG	BamHI-
	Reverse	CCCGCTCGAG-ACGCCGACGTTTTTTATTAA	NdeI
767		CGCGGATCCCATATG-CTGACGGAAGGGGAAG	XhoI
707	Torward	CGCGGATCCCATATG-CTGACGGAAGGGGAAG	BamHI-
	Reverse	CCCGCTCGAG-TTTCTGTACAGCAGGGG	NdeI *
768		CGCGGATCCCATATG-GCCCCGCAAAAACCCG	XhoI
, 00	1 OI Wald	COCOCATATO-OCCCCOCAAAACCCCG	BamHI-
	Reverse	CCCGCTCGAG-TTTCATCCCTTTTTTGAGC	NdeI
770		CGCGGATCCCATATG-TGCGGCAGCGGCGAA	XhoI
	1 01	-reedeAdeddedAA	BamHI-
	Reverse	CCCGCTCGAG-GCGTTTGTCGAGATTTTC	NdeI XhoI
771		CGCGGATCCCATATG-TCCGTATATCGCACCTTC	Anoi BamHI-
		TOOGIMIATOOCACCITC	NdeI
	Reverse	CCCGCTCGAG-CGGTTCTTTAGGTTTGAG	Ndel XhoI
772		CGCGGATCCCATATG-TTTGCGGCGTTGGTGG	BamHI-
			Ndel
	Reverse	CCCGCTCGAG-CAATGCCGACATCAAACG	XhoI
774		CGCGGATCCCATATG-TCCGTTTCACCCGTTCC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-TCGTTTGCGCACGGCT	XhoI
790		CGCGGATCCCATATG-GCAAGAAGGTCAAAAAC	BamHI-
			Ndel
			11001

Reverse CCCGCTCGAG-GGCGTTGTTCGGATTTCG  Porward CGCGGATCCCATATG-CCGTCTGAAATGCCG  Reverse CCCGCTCGAG-ATATGGAAAAGTCTGTTGTC  Ndel  Reverse CCCGCTCGAG-AAAATGGAAAAATGCAGG  Porward CGCGGATCCCATATG-CCGATTTTTCGATG  Reverse CCCGCTCGAG-AAAATGGAACAATACCAGG  Porward CCGGAATCTACATATG-TTGCACTTTCAAAGGATAATC  Reverse CCCGCTCGAG-AAAAATGTACAATGGCGTAC  Reverse CCCGCTCGAG-AAAAATGTACAATGGCGTAC  Reverse CCCGCTCGAG-AAAAAATGTACAATGGCGTAC  Reverse CCCGCTCGAG-AAAAAATGTACAATGGCGTAC  Reverse CCCGCTCGAG-AAAAAATGTACAATGGCGTAC  Reverse CCCGCTCGAG-GAAACATGTAATTCAAGTTGAA  Porward AAAAAAGGTACC-ATGATGCAGCACAATCGTTTC  Reverse AAACTGCAG-TTAATATCGATAGGTTAATATG  Porward AAAAAAAGAATTC-CGGCTCGGCATTGTGCAGATGTTGCA  Reverse AAACTGCAG-TTAATATCGATAGGTTAATATG  Porward CGCGGATCCCATATG-AACAAAATATACCGCATC  Reverse CCCGCTCGAG-CCACTGATAACCGACAGAT  Ndel  Reverse CCCGCTCGAG-CCACTGATAACCGACAGAT  Ndel  Reverse CCCGCTCGAG-CCACTGATAACCGACAGAT  Ndel  Reverse CCCGCTCGAG-ACGCCACTGATAACCGACAGAT  Reverse CCCGCTCGAG-CCACTGATAACCGACAGAT  Reverse CCCGCTCGAG-TCGATAACCGACAGAT  Reverse CCCGCTCGAG-TCGATAACCGACAGAT  Reverse CCCGCTCGAG-TCGACTGCAGCG  Porward AAAAGAATTC-GCAGATTGTAGAGCGTTAATAAAAATAC  Reverse CCCGCTCGAG-TCGATAACCGACAGGC  Reverse CCCGCTCGAG-TCGACTGCAGCG  Porward AAAAACTGCAG-TTAATATGGTTTTTTTTCGTTCG  Porward AAAAAACTGCAG-TTAATATGGTTTTTTTTCGTTCG  Porward AAAAAACTGCAG-TCGACTCGCCGCGACCTCTGCCGAGCG  Reverse AAACTGCAG-TCAGCGATCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC		Reverse CCCCCTCGAC CCCCTTCTTGGGATTTGG	
Reverse CCCGCTCGAG-ATATGGAAAAGTCTGTTGTC  Pornard CGCGGATCCCATATG-CCCGATTTTTCGATG  Reverse CCCGCTCGAG-AAAATGGAACAATACCAGG  Pornard CGCGAATTCTACATATG-TTGCACTTTCAAAGGATAATC  Reverse CCCGCTCGAG-AAAAATGTACAATGGCGTAC  Reverse CCCGCTCGAG-AAAAATGTACAATGGCGTAC  Reverse CCCGCTCGAG-GAAAACTGTAATTCAAGTTGAA  Pornard AAAAAAGGTTAC-CAGCGTCAGCACAATCGTTC  Reverse AACTGCAG-TTAATATCGATAGGTTAATG  Pornard AAAAAAAGAATTC-CGGCTCGGCATGTGCAGCAGCAT  Reverse CCCGCTCGAG-AAAAATGTACAATGGTTAATG  Pornard AAAAAAAGAATATAGATAGGTAATATG  Pornard CGCGGATCCCATATG-AACAAAATATACGATAGGTAATG  Pornard CGCGGATCCCATATG-AACAAAATATACGATAGGTAATAG  Pornard CGCGGATCCCATATG-AACAAAATATACCATCAGTAGAGTAG	900	Forward CGCGGATCCCATATC GGGTGTGATATCG	XhoI
Reverse CCCGCTCGAG-ATATGGAAAAGTCTGTTGTC  Reverse CCCGCTCGAG-AAAATGGAACAATACCAGG  Porward CGCGGATCCCATATG-CCCGATTTTCGATG  Reverse CCCGCTCGAG-AAAAATGAACATGCAGG  Reverse CCCGCTCGAG-AAAAATGTACAATGGCGTAC  Possible Forward CGCGAATTCTAGCTAGC-CAGCGTCAGCAGCACAT  Reverse CCCGCTCGAG-AAAAATGTACAATGGCGTAC  Reverse CCCGCTCGAG-GAAACTGTAATTCAAGTTGAA  Possible Forward AAAAAAGGTACC-ATGATGCAGCACAATCGTTTC  Reverse AAACTGCAG-TTAATATCGATAGGTTAATTG  Possible Forward AAAAAAGAATC-CGGCTCGGCATTGTGCAGCAGTTGCA  Reverse AAACTGCAG-TTAATATCGATAGGTTAATTG  Possible Forward CGCGGATCCCATATG-AACAAAATATACCGCATC  Reverse CCCGCTCGAG-CCACTGATAACCGACAGT  Reverse CCCGCTCGAG-CCACTGATAACCGACAGT  Reverse CCCGCTCGAG-CCACTGATAACCGACAGT  Reverse CCCGCTCGAG-ACGCCACTGCCAGCG  Possible Forward CGCGGATCCCATATG-TGCGCGTGGAAACTTAT  Reverse CCCGCTCGAG-ACGCCACTGCCAGCG  Possible Forward CGCGGATCCCATATG-TGCGCGTGGAAACTTAT  Reverse CCCGCTCGAG-TCAATATGATGTTTT  Reverse CCCGCTCGAG-TCGATATGAGTTATATATATATATATATATATATATATAT	100	Forward CGCGGATCCCATATG-CCGTCTGAAATGCCG	BamHI-
901   Forward   CGCGGATCCCATATG-CCCGATTTTTCGATG   BamHI-Ndel		Persona CCCCCTCCAC ATATOCAAAA	NdeI
Reverse CCCGCTCGAG-AAAATGGAACAATACCAGG  Porward. CCGGAATTCTACATATG-TTGCACTTTCAAAGGATAATC  Reverse CCCGCTCGAG-AAAAATGTACAATGGCGTAC  Reverse CCCGCTCGAG-GAAACTGTACC-CAGCGTCAGCAGCACAT  Reverse CCCGCTCGAG-GAAACTGTAATTCAAGTTGAA  904 Forward AAAAAAGGTACC-ATGATGCAGCACAATCGTTTC  Reverse AAACTGCAG-TTAATATCGATAGGTTAATATG  905 Forward AAAAAAGGATTC-CGGCTCGGCATTGTGCAGCAGATGTTGCA  Reverse AAACTGCAG-TTAATATCGATAGGTTAATATG  906 Forward AAAAAAGAATTC-CGGCTCGGCATTGTGCAGATGTTGCA  Reverse CCCGCTCGAG-CACATGATACCGACAGAT  Reverse CCCGCTCGAG-CCACTGATAACACGACAGAT  Pst 1  907 Forward CGCGGATCCCATATG-AACAAAATATACCGCATC  Reverse CCCGCTCGAG-CCACTGATAACCGACAGAT  Reverse CCCGCTCGAG-ACGCCACTGCCAGCG  Post 1  Reverse CCCGCTCGAG-ACGCCACTGCCAGCG  Reverse AAACTGCAG-TTAATATGGTTTTGTCGTTCG  Pst 1  Reverse CCCGCTCGAG-TCGATTTGAAACTTTGGTTTT  Ndel  Reverse CCCGCTCGAG-TCGGTTTTGAAACTTTGGTTTT  Ndel  Reverse CCCGCTCGAG-TCGGTTTTGAAACTTTGGTTTT  Pst 1  Reverse AAACTGCAG-TCAGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC	001	Forward CCCCCATCCCATATO COATATO	XhoI
Reverse CCGCTCGAG-AAAATGGAACAATACCAGG  Potward. CCGGAATTCTACATATG-TTGCACTTTCAAAGGATAATC Reverse CCCGCTCGAG-AAAAATGTACAATGGCGTAC  Reverse CCCGCTCGAG-AAAAATGTACAATGGCGTAC  Reverse CCCGCTCGAG-GAAAAATGTACAATGGCGTAC  Reverse CCCGCTCGAG-GAAACTGTAATTCAAGTTGAA  Potward AAAAAAGGTACC-ATGATGCAGCACAATCGTTTC Reverse AAACTGCAG-TTAATATGGATAGGTTAATTG  Reverse AAACTGCAG-TTAATATCGATAGGTTAATTG  Potward AAAAAAGAATTC-CGGCTCGGCATTGTGCAGAGTAGTTGCA Reverse AAACTGCAG-CAATGATAACCGACACAATCGTTGCA Reverse CCCGCTCGAG-CCACTGATAACCGACACATC  Reverse CCCGCTCGAG-CCACTGATAACCGACACAT  Reverse CCCGCTCGAG-CCACTGATAACCGACACAT  Reverse CCCGCTCGAG-ACGCCACTGCCACGG  Potward CGCGGATCCCATATG-GGCGCGCAACGTGAG  Reverse CCCGCTCGAG-ACGCCACTGCAGCG  Potward AAAGAATTC-GCAGAGGTTAGTAGGCGTTAATAAAAATAC Reverse AAACTGCAG-TCAGTTAGTAGGCGTTAATAAAAATAC Reverse AAACTGCAG-TCGATATG-TGCCGCGGGAACCTTAT  Reverse CCCGCTCGAG-CCCACTGCCAGCG  Potward CGCGGATCCCATATG-TGCGCGTGGGAAACTTAT  Reverse CCCGCTCGAG-CCGCATTTGAAACTTTGGTTTT  Reverse CCCGCTCGAG-TCGGTTTTGAAACTTTGGTTTT  Reverse CACGGTCCAATTG-CGCTTGCCGGGGACCTCTTT  Potward AAAGAATTC-GCATTTGCCGGGGACTCTCTCCCGAGCG Reverse AAACTGCAG-TCAGCGACTCAACTGCTCTTT  Potward AAAGAATTC-GCTTTCCGCGTGGCCGGCGTGC Reverse AAAACTGCAG-TCAGCTTATTCGCCGGCGCTTTTTCCCC  Reverse AAACTGCAG-TTACAGTCCGTCCACGCCTTTTCCCC CAAATCCGTCAAAACGCCACTCAAGTATTGAG  Reverse CCCGCTCGAG-ACGTCGACTTATCGCCGCCGCCCCCC  Reverse CAAAAACTGCAG-TTACAGTCCGTCCACGCCTTTCCCC  Pot I  Reverse CCCGCTCGAG-ACGTTGTCCACGCCTTTCCC  Pot I  Reverse CCCGCTCGAG-ACGTTGTCCACGCCTTCCC  Remillion AAACACTCCATATG-TGCCGCGCACGCTACACGC  Reverse CCCGCTCGAG-TCTTGAAAATATAGGTTTCAACCG  Reverse CCCGCTCGAG-TCTTGAAAATATAGGTTTCAACCG  Reverse CCCGCTCGAG-TCTTGAAAATATAGGTTTCAACCG  Reverse CCCGCTCGAG-TCTTTGA	901	Forward CGCGGATCCCATATG-CCCGATTTTTCGATG	BamHI-
Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot   Pot		Payarra CCCCCTCCAC AAAARCCA	NdeI
Reverse CCCGCTCGAG-AAAAATGTACAATGGCGTAC  Reverse CCCGCTCGAG-GAACTGTAATTCAAGTTGAA  904 Forward AAAAAAGAATTC-CGGCTCGGCACACATCTTC  Reverse AAACTGCAG-TTAATATCGATAGGTTATATG  905 Forward AAAAAAGAATTC-CGCGCCGCATTGTCAGATGTGCA  Reverse AAACTGCAG-TTAATATCGATAGGTTATATG  906 Forward AAAAAAGAATTC-CGGCTCGGCATTGTCAGATGTTGCA  Reverse AAACTGCAG-TTAATATCGATAGGTTATATG  907 Forward CGCGGATCCCATATG-AACAAAATATACCGCATC  Reverse CCCGCTCGAG-CCACTGATAACCGACAGAT  908 Forward CGCGGATCCCATATG-GGCGCGCAACGTGAG  809 Forward AAAGAATTC-GCAGAGTTAGTAGGTTAATACCGACAGAT  909 Forward CGCGGATCCCATATG-TGCCGGCGCAACGTGAG  800 Forward AAAGAATTC-GCAGAGTTAGTAGGCGTTAATAAAAATAC  800 Forward CGCGGATCCCATATG-TGCGCGGGAAACTTAT  800 Forward CGCGGATCCCATATG-TGCGCGGGAAACTTAT  800 Forward CGCGGATCCCATATG-TGCGCGTGGAAACTTAT  800 Forward CGCGGATCCCATATG-TGCGCGTGGAAACTTAT  800 Forward CGCGGATCCCATATG-TGCGCGTGGAAACTTAT  800 Forward CGCGGATCCCATATG-TGCGCGTGGAAACTTAT  800 Forward CGCGGATCCCATATG-TGCGCGTGGGAAACTTAT  800 Forward CGCGGATCCCATATG-TGCGCGTGGCAGCG  800 Forward AAAGAATTC-GCATTTGCCGGCGGACTCTTTT  801 Forward AAAGAATTC-GCATTTGCCGGCGGACTCTTTT  910 Forward AAAGAATTC-GCATTTGCCGGCGGACTCTTTT  911 Forward AAAGAATTC-GCATTTGCGCGGCGGCGTGC  801 Forward AAAAAACTGCAG-TCAACGCATCAAGTATTGAG  802 FORWARD AAAAAACTGCAG-TCAACACCCCCCCC  CAAATCCGCAG-TTACAGTCCTCCACGCTTTTCCCC  912 Forward CGCGGATCCCATATG-GAAACCCCACTCAAGTATTGAG  803 Forward CGCGGATCCCATATG-GAAACCCCACTCAAGTATTGAG  804 Forward CGCGGATCCCATATG-GAAACCCCACTCAAGTATTGAG  805 Forward CGCGGGATCCCATATG-GAAACCCCACTCAAGTATTGAG  806 FORWARD CGCGGATCCCATATG-GAAACCCCACTCAAGTATTGAG  807 FORWARD CGCGGATCCCATATG-GAAACCCCACTCAAGTATTGAG  808 FORWARD CGCGGATCCCATATG-GAAACCCGCCCCGC  807 FSt I  808 FORWARD CGCGGATCCCATATG-GAAACCCGCCCCGC  808 FI  809 FORWARD CGCGGGATCCCATATG-GAAACCCGCCCCGC  807 FST I  807 FORWARD CGCGGGATCCCATATG-GAAACCCGCCCCGC  807 FST I  807 FORWARD CGCGGGATCCCATATG-GAAACCCGCCCCGC  807 FST I  808 FORWARD CCGGGATCCCATATG-GAAACCCGCCCGCCGCCGCGCGAACCGCTCAACCCCCCCC	002	Reverse CCCGCTCGAG-AAAATGGAACAATACCAGG	XhoI
Reverse CCCGCTCGAG-AAAAATGTACAATGGCGTAC  Poil Forward CCGGAATTCTAGCTAGC-CAGCGTCAGCAGCACAT  Reverse CCCGCTCGAG-GAAACTGTAATTCAAGTTGAA  Poil Forward AAAAAAGGTACC-ATGATGCAGCACACATCGTTTC Reverse AAACTGCAG-TTAATATCGATAGGTTATATG  Poil Forward AAAAAAGAATTC-CGGCTCGGCATTGTGCAGATGTTGCA Reverse AAACTGCAG-TTAATATCGATAGGTTATATG  Poil Forward CCGCGTCGAG-CCACTGATAACCGACACATCGTTC Reverse CCCGCTCGAG-CCACTGATAACCGACATC  Reverse CCCGCTCGAG-CCACTGATAACCGACAGAT  Poil Forward CGCGGATCCCATATG-GGCGCGCAACCGTGAG  Reverse CCCGCTCGAG-ACGCCACTGCCAGCG Reverse AAACTGCAG-TTAATATGGTTTTGTCGTTCG  Poil Forward CGCGGATCCCATATG-TGCGCGGCATCTATAAAAAATAC Reverse CCCGCTCGAG-TCGGTTTTGAAACTTTGTTT  Reverse CCCGCTCGAG-TCGGTTTTGAAACTTTGTTT  Reverse CCCGCTCGAG-TCGGTTTTGAAACTTTGTTT  Poil Forward AAAGAATTC-GCATTGCCGGCGACTCTTT  Poil Forward AAAGAATTC-GCATTTGCCGTGCGCGGCGTGC Reverse AAACTGCAG-TCAGCGATCGAGCTGCCGAGCG Reverse AAACTGCAG-TCAGCGATCGAGCTGCCGCGCGTGC Reverse AAAACTGCAG-TCAGCGTCGCGCGCGTGC Reverse AAAAAAACTGCAG-TCAGCGTCGCGCGCGTTTTTCCGC  Poil Forward AAAAAAACTGCAG-TCAGCGTCCACGCCTTTTCCGC  Poil Forward CGCGGATCCCATATG-GAAACCCGCCCCGC Reverse AAAAAAACTGCAG-TCAGCGTTCAACGCCCCCGC  Poil Forward CGCGGATCCCATATG-GAAACCCGCCCCCCC  Reverse CCCGCTCGAG-AGGTTGTTTCCAGCTGTCCACGCCTTTTCCCC  Poil Reverse CCCGCTCGAG-AGGTTGTTCCAGCTTCCACGCCTTTTCCCC  Poil Reverse CCCGCTCGAG-AGGTTGTTCCAGCTTTCCAGCTTTCCCC  Poil Reverse CCCGCTCGAG-AGGTTGTTCCAGCTTCCACGCCTTTCCCC  Poil Reverse CCCGCTCGAG-AGGTTGTTTCCAGCTTCCAGCCCTTTCCCC  Poil Reverse CCCGCTCGAG-AGGTTGTTCCAGCTTTCCAGCTTTCCAGCTTTCCAGCTTTCCAGCTTTCCAGCTTTCCAGCTTTCCAGCTTTCCAGCTTTTCCAGCTTTCCAGCTTTTCCAGCTTTCCAGGTTG  Reverse CCCGCTCGAG-TTTGAAAACCCGCCCCCCCCCCCCCCCCC	902	Folward. CCGGAATTCTACATATG-TTGCACTTTCAAAGGATAATC	EcoRI-
Forward CCGGAATTCTAGCTAGC-CAGCGTCAGCAGCACAT  Reverse CCCGCTCGAG-GAAACTGTAATTCAAGTTGAA  904 Forward AAAAAAGGTACC-ATGATGCAGCACAATCGTTTC Reverse AAACTGCAG-TTAATATCGATAGGTTATATG  904a Forward AAAAAAGAATTC-CGGCTCGGCATTGTGCAGAGTGTTGCA Reverse AAACTGCAG-TTAATATCGATAGGTTATATG  905 Forward CGCGGATCCCATATG-AACAAAATATACGCATC  8 BamHI- Reverse CCCGCTCGAG-CCACTGATAACCGACAGAT  907 Forward CGCGGATCCCATATG-GGCGCGCAACGTGAG  8 BamHI- Reverse CCCGCTCGAG-ACGCCACTGCCAGCG  908 Forward AAAGAATTC-GCAGAGTTAGTGTTCCG  909 Forward CGCGGATCCCATATG-TGCGCGCGCGTAATAAAAAATAC Reverse AAACTGCAG-TTAATATGGTTTTTTCTCGTTCG  909 Forward AAAGAATTC-GCAGTTGTGAAACTTTGGTTTT  Reverse CCCGCTCGAG-TCGGTTTTGAAACTTTGGTTTT  Ndel Xhol Reverse AAACTGCAG-TCAGCTTTGCAGCGAGCG Reverse AAACTGCAG-TCAGCACTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC			NdeI
Reverse CCCGCTCGAG-GAAACTGTAATTCAAGTTGAA  Nhel  Nhol  Potation of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property of the property	003	Forward CCCCAATTCTACCTACCTAC	XhoI
Reverse CCCGCTCGAG-ACACTGTAATTCAAGTTGAA  904 Forward AAAAAAGGTACC-ATGATGCAGCACAAATCGTTTC Reverse AAACTGCAG-TTAATATCGATAGGTTAATATG  904a Forward AAAAAAGAATTC-CGGCTCGGCATTGTGCAGATGTTGCA Reverse AAACTGCAG-TTAATATCGATAGGTTATATG  905 Forward CGCGGATCCCATATG-AACAAAATATACCGCATC  88mHI- Ndel Ndel Ndel Ndel Ndel Ndel Ndel Ndel	903	Folward CCGGAATICTAGCTAGC-CAGCGTCAGCACAT	EcoRI-
Forward AAAAAAGGTACC-ATGATGCAGCACAATCGTTTC Reverse AAACTGCAG-TTAATATCGATAGGTTATATG Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I Pst I P		Perverse CCCCCCCAC CAAACTCTAATTC	NheI
Reverse AAACTGCAG-TTAATATCGATAGGTTATATG  904a Forward AAAAAAAGAATTC-CGGCTCGGCATTGTGCAGATGTTGCA Reverse AAACTGCAG-TTAATATCGATAGGTTATATG  905 Forward CGCGGATCCCATATG-AACAAAATATACCGCATC  Reverse CCCGCTCGAG-CCACTGATAACCGACAGAT  907 Forward CGCGGATCCCATATG-GGCGCGCAACGTGAG  8	904	Forward AAAAACCTACC ATCATTCAAGTTGAA	XhoI
Forward AAAAAAGAATTC-CGGCTCGGCATTGTGCAGATGTTGCA Reverse AAACTGCAG-TTAATATCGATAGGTTATATG  905 Forward CGCGGATCCCATATG-AACAAAATATACCGCATC  Reverse CCCGCTCGAG-CCACTGATAACCGACAGAT  907 Forward CGCGGATCCCATATG-GGCGCGCAACGTGAG  908 Forward AAAGAATTC-GCAGAGTTAGTAGGCGTTAATAAAAAATAC Reverse CCCGCTCGAG-ACGCCACTGCCAGCG  909 Forward AAAGAATTC-GCAGAGTTAGTAGGCGTTAATAAAAAATAC Reverse AAACTGCAG-TTAATATGGTTTTGTCGTTCG  909 Forward CGCGGATCCCATATG-TGCGCGGGAAACTTAT  82  82  83  84  84  84  85  85  86  80  81  81  82  83  84  84  84  84  84  84  84  84  84	1 304	Powers AAAAAAGGTACC-ATGATGCAGCACAATCGTTTC	Kpn I
Reverse AAACTGCAG-TTAATATCGATAGGTTATATG  905 Forward CGCGGATCCCATATG-AACAAAATATACCGCATC  Reverse CCCGCTCGAG-CCACTGATAACCGACAGAT  907 Forward CGCGGATCCCATATG-GGCGCGCAACGTGAG  Reverse CCCGCTCGAG-ACGCCACTGCCAGCG  908 Forward AAAGAATTC-GCAGAGTTAGTAGGCGTTAATAAAAAATAC  Reverse AAACTGCAG-TTAATATGGTTTTGTCGTTCG  909 Forward CGCGGGATCCCATATG-TGCGCGTGGAAACTTAT  Reverse CCCGCTCGAG-TCGGTTTTGAAACTTTGGTTTT  Ndel  Reverse CCCGCTCGAG-TCGGTTTTGAAACTTTGGTTTT  Porward AAAGAATTC-GCATTTGCCGGCGAGCG  Reverse AAACTGCAG-TCAGCGATCGAGTGCC  Reverse AAACTGCAG-TCAGCGATCGAGTGCC  Reverse AAAAAACTGCAG-GTCGACTTATTCGGCGGCGTTTTTCCGC  Pst I  Forward AAAAAAACTGCAG-GTCGACTTATTCGGCGCTTTTCCGC  Pst I  Eco RI  Reverse CCCGCTCGAG-TCAGCCACTCAAGTATTGAG  Reverse AAAAAACTGCAG-TTACAGTCCGTCCACGCTTTCCC  Pst I  Eco RI  Pst I  Forward CGCGGATCCCATATG-GAAACCCCCCCCC  Pst I  Reverse CCCGCTCGAG-TGTTTCCGCGTCCACGCTTTCCCC  Pst I  Eco RI  Reverse CCCGCTCGAG-AGGTTGTTTCCAGGTTG  Ndel  Ndel  Ndel  Reverse CCCGCTCGAG-TTTGAAAACTCCGCCCGCAGAA  BamHI-  Ndel  Ndel  Ndel  Ndel  Ndel  Ndel  Ndel  Ndel  Ndel  Ndel  Ndel  Ndel  Ndel  Ndel  Ndel  Nol  Forward AAAGAATTC-GACAGAATTGAGG  Proward AAAGAATTC-GACAGAATTGAGG  Reverse CCCGCTCGAG-TTTGAAAATTAGGTATCAGG  Pst I  Reverse CCCGCTCGAG-TTTGAAAACTGCCGCCCCGC  Reverse CCCGCTCGAG-TTTGAAACCCGCCCCGC  Reverse CCCGCTCGAG-TTTGAAAACTGCGCGAACCCG  Reverse CCCGCTCGAG-TTTGAAAATTAGGTATCAGG  Pst I  Reverse CCCGCTCGAG-TTTGAAAACTGCCGCCCCGC  Reverse CCCGCTCGAG-TTTGAAAACTGCGCGCAGCGGAA  Reverse CCCGCTCGAG-TTTGAAAATTAGGTATCAGG  Pst I  Reverse CCCGCTCGAG-TTTGAAACCCGCCCCCCCCCCCCCCCCCC	004-	Reverse AAACTGCAG-TTAATATCGATAGGTTATATG	Pst I
Forward CGCGGATCCCATATG-AACAAAATATACCGCATC  Reverse CCCGCTCGAG-CCACTGATAACCGACAGAT  Porward CGCGGATCCCATATG-GGCGCGCAACGTGAG  Reverse CCCGCTCGAG-ACGCCACTGCCAGCG  Reverse CCCGCTCGAG-ACGCCACTGCCAGCG  Porward AAAGAATTC-GCAGAGTTAGTAGGCGTTAATAAAAAATAC Reverse AAACTGCAG-TTAATATGGTTTTGTCGTTCG  Porward CGCGGATCCCATATG-TGCGCGTGGAAACCTTAT  Reverse CCCGCTCGAG-TCGGTTTTGAAACTTTGGTTTT  Reverse CCCGCTCGAG-TCGGTTTTGAAACTTTGGTTTT  Porward AAAGAATTC-GCATTTGCCGGCGAGCG Reverse AAACTGCAG-TCAGCGATCGAGCTGCTCTTT  Porward AAAGAATTC-GCTTTCCGCGTGGCCGGCGTGC Reverse AAAAAACTGCAG-GTCGACTTATTCGGCGGCTTTTTCCGC  Post I  Eco RI  Porward AAAAAACTGCAG-TTACAGTCCGTCCACGCTTTCCCC  Reverse AAAAAACTGCAG-TTACAGTCCGTCCACGCTTTCCCC  Reverse CCCGCTCGAG-ACGCTCAAGTATTGAG Reverse CCCGCTCGAG-AGGTTGTTCCAGCTTCCCCCGCCCCCCC  Post I  Eco RI  Porward CGCGGATCCCATATG-GAAACCCCCCCCCC  Post I  Reverse CCCGCTCGAG-AGGTTGTTCCAGGTTG  Ndel  Ndel  Ndel  Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG  Porward AAAGAATTC-GCCGGCAGGCGGAA  Reverse CCCCGCTCGAG-TTTGAAAATATAGGTATCAGG  Porward AAAGAATTC-GACAGAATCGGCGAAGCACG  Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG  Porward AAAGAATTC-GACAGAATCGGCGAAGCACG  Reverse AAACTGCAG-TTTGAAAATTAGGTATCAGG  Porward AAAGAATTC-GACAGAATCGGCGAATTTGGAAGCACG  Reverse AAACTGCAG-CTATATG-GCCGGCAGGACCGCCCCCCCCCC	904a	Poward AAAAAAGAATTC-CGGCTCGGCATTGTGCAGATGTTGCA	Eco RI
Reverse CCCGCTCGAG-CCACTGATAACCGACAGAT  907 Forward CGCGGATCCCATATG-GGCGCGCAACGTGAG  Reverse CCCGCTCGAG-ACGCCACTGCCAGCG  908 Forward AAAGAATTC-GCAGAGTTAGTAGGCGTTAATAAAAAATAC Reverse AAACTGCAG-TTAATATGGTTTTGTCGTTCG  909 Forward CGCGGATCCCATATG-TGCGCGTGGAAACTTAT  Reverse CCCGCTCGAG-TCGGTTTTGAAACTTTGGTTTT  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  Pst I  BamHI-  NdeI  NdeI  NdeI  Shoil  Forward AAAGAATTC-GCATTTGCGGCGTGGAAACTTAT  Reverse AAACTGCAG-TCGGTTTTGAAACTTTGGTTTT  Pst I  Forward AAAGAATTC-GCATTTGCCGGCGAGCG Reverse AAAACTGCAG-TCAGCGATCGACTGCTCTTT  Pst I  Forward AAAAAACTGCAG-GTCGACTTATTCGGCGGCGTTTTTCCGC  Pst I  Forward AAAAAACTGCAG-TTACAGTCCGTCCACGCCTTTCGC  Reverse AAAAAACTGCAG-TTACAGTCCGTCCACGCCTTTCGC  Pst I  Reverse CCCGCTCGAG-AGGTTGTTCCAGGTTG  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  Nde	005	Reverse AAACTGCAG-TTAATATCGATAGGTTATATG	Pst I
Reverse CCCGCTCGAG-CCACTGATAACCGACAGAT  907 Forward CGCGGATCCCATATG-GGCGCGCAACGTGAG  Reverse CCCGCTCGAG-ACGCCACTGCCAGCG  908 Forward AAAGAATTC-GCAGAGTTAGTAGGCGTTAATAAAAAATAC  Reverse AAACTGCAG-TTAATATGGTTTTGTCGTTCG  909 Forward CGCGGATCCCATATG-TGCGCGTGGAAACTTAT  Reverse CCCGCTCGAG-TCGGTTTTGAAACTTTGGTTTT  Pst I  Poward AAAGAATTC-GCATTTGCCGGCGAGCG  Reverse AAACTGCAG-TCAGCTTCGCGAGCTCTTT  911 Forward AAAGAATTC-GCTTTCCGCGTGGCCGGCGTGC  Reverse AAAACTGCAG-GTCGACTTATTCCGCCGCGCGTTTTCCGC  912 Forward AAAAAACTGCAG-GTCGACTTATTCGGCGGCTTTTTCCGC  Pst I  Pst I  Eco RI  Reverse AAAAAACTGCAG-GTCGACTTATTCGGCGGCTTTTTCCGC  Pst I  Eco RI  Reverse AAAAAACTGCAG-TTACAGTCCGTCCACGCCTTTCGC  Pst I  Forward CGCGGATCCCATATG-GAAACCCGCCCCGC  Reverse CCCGCTCGAG-AGGTTGTTCCAGGTTG  Pst I  Reverse CCCGCTCGAG-AGGTTGTTCCAGGTTG  Ndel  Xhol  Reverse CCCGCTCGAG-TTGAAAACCCGCCCCGC  BamHI-  Ndel  Ndel  Xhol  Pst I  Eco RI  Reverse CCCGCTCGAG-AGGTTGTTCCAGGTTG  AAAAAACTGCAG-TTACAGTCCGTCACGCCTTTCGC  Pst I  Reverse CCCGCTCGAG-TGGTTTCCAGGTTG  Ndel  Xhol  Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG  Pst I  Reverse CCCGCTCGAG-TTTGAAAACCCGCCCCGC  Reverse CCCGCTCGAG-TTTGAAAACCCGCCCCCGC  Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG  Pst I  Reverse CCCGCTCGAG-TTTGAAAACCCGCCCCCGC  Reverse CCCGCTCGAG-TTTGAAAACCCGCCCCCGC  Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG  Pst I  Reverse CCCGCTCGAG-TTTGAAAACCCGCCCCCGC  Reverse CCCGCTCGAG-TTTGAAAACCCGCCCCCCCCCCCCCCCCC	905	Forward CGCGGATCCCATATG-AACAAAATATACCGCATC	BamHI-
Forward CGCGGATCCCATATG-GGCGCGCAACGTGAG  Reverse CCCGCTCGAG-ACGCCACTGCCAGCG  908 Forward AAAGAATTC-GCAGAGTTAGTAGGCGTTAATAAAAAATAC Reverse AAACTGCAG-TTAATATGGTTTTGTCGTTCG  909 Forward CGCGGATCCCATATG-TGCGCGTGGGAAACTTAT  Reverse CCCGCTCGAG-TCGGTTTTGAAACTTTGTTTT  Porward AAAGAATTC-GCATTTGCCGGCGAGCG Reverse AAACTGCAG-TCAGCGATCGACTCTTT  911 Forward AAAGAATTC-GCTTTCCGCGTGGCAGCG Reverse AAAAAACTGCAG-GTCGACTTATTCGGCGGCGTGC Reverse AAAAAACTGCAG-GTCGACTTATTCGGCGGCGTTTTCCGC  912 Forward AAAAAAACTGCAG-TTACAGTCCGTCCACGCCTTTCCGC Reverse AAAAAACTGCAG-TTACAGTCCGTCCACGCCTTTCCGC  913 Forward CGCGGATCCCATATG-GAAACCCGCCCCGC  Reverse CCCGCTCGAG-AGGTTGTTCCAGGTTG  Reverse CCCGCTCGAG-AGGTTGTTCCAGGTTG  Poward CGCGGATCCCATATG-GAAACCCGCCCCGC  Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG  Poward CGCGGATCCCATATG-TGCCGGCAGCGGAA  Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG  Poward AAAGAATTC-GACAGAATCGGCGATTTGGAAGCACG  Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG  Poward AAAGAATTC-GACAGAATCGGCGATTTGGAAGCACG  Reverse AAACTGCAG-CTATATGCGCCGCAGGACGCCTCAACGC  Poward AAAGAATTC-GACAGAATCGGCGATTTTGGAAGCACG  Reverse AAACTGCAG-CTATATGGGCCGGCAGGACGCTCAACGC  Reverse AAACTGCAG-CTATATGCGCCGCAGGACGCTCAACGC  Reverse AAACTGCAG-CTATATGCGCCGGCAGGACGCTCAACGC  Reverse AAACTGCAG-CTATATGCGCCGGCAGGACGCTCAACGC  Reverse AAACTGCAG-CTATATGCGCCGGCAGGACGCTCAACGC  Reverse AAACTGCAG-CTATATGCGCCGCAGGACGCTCAACGC  Reverse AAACTGCAG-CTATATGCGCCGGCAGGACGCTCAACGC  Reverse AAACTGCAG-CTATATGCGCCGGCAGGACGCTCAACGCC  Reverse AAACTGCAG-CTATATGCGCCGGCAGGACGCTCAACGCC  Reverse AAACTGCAG-CTATATGCGCCGGCAGGACGCTCAACGCC  Reverse AAACTGCAG-CTATATGCGCCGGCAGGCACGCTCAACGCC  Reverse AAACTGCAG-CTATATGCGCCGGCAGGCACGCTCAACGCC  Reverse AAACTGCAG-CTATATGCGCCGGCAGGCACGCTCAACGCC  Reverse AAACTGCAG-CTATATGCGCCGGCAGGCGCTCAACGCC  Reverse AAACTGCAG-CTATATGCGCCGGCACGCTCAACGCC  Reverse AAACTGCAG-CTATATGCGCCGGCACGCTCAACGCC  Reverse AAACTGCAG-CTATATATGCGCCGGCACGCTCAACGCC  Reverse AAACTGCACTATATGCGCCGCACGCTCAACGCC  Reverse AAACTGCACTATATGCGCCGCCTCAACGCC  Reverse AAACTGCACTATATGCGCCGCCCTCAACGCC  Reverse AAACTGCACTATATGCGCCG		Personal CCCCCTCC AC CC AC CC	
Reverse CCCGCTCGAG-ACGCCACTGCCAGCG  908 Forward AAAGAATTC-GCAGAGTTAGTAGGCGTTAATAAAAATAC Reverse AAACTGCAG-TTAATATGGTTTTGTCGTTCG  909 Forward CGCGGATCCCATATG-TGCGCGTGGGAAACTTAT  Reverse CCCGCTCGAG-TCGGTTTTGAAACTTTGGTTTT  Porward AAAGAATTC-GCATTTGCCGGCGAGCG Reverse AAACTGCAG-TCAGCGATCGACTCTTT  911 Forward AAAGAATTC-GCTTTCCGCGTGGCCGTGC Reverse AAAAAAACTGCAG-GTCGACTTATTCGGCGGTGC Reverse AAAAAAACTGCAG-GTCGACTTATTCGGCGTTTT  912 Forward AAAAAAACTGCAG-TTACAGTCGCGCGCGCTTTT  Porward AAAAAAACTGCAG-TTACAGTCCGCCGCGCTTTCCGC Pst I  Eco RI  Reverse AAAAAAACTGCAG-TTACAGTCCGTCCACGCCTTTCGC Pst I  Forward CGCGGATCCCATATG-GAAACCCGCCCCGC  Porward CGCGGATCCCATATG-GAAACCCGCCCCGC  Porward CGCGGATCCCATATG-TGCCGGCGAGCGAA  Reverse CCCGCTCGAG-AGGTTGTTTCCAGGTTG  Porward CGCGGATCCCATATG-TGCCGGCAGCGGAA  Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG Porward AAAGAATTC-GACAGAATCGGCGAACGCCTCAACCCG  Porward CGCGCTCGAG-TTTGAAAATATAGGTATCAGG Porward AAAGAATTC-GACAGAATCGGCGAACGCCGCCCGC  Porward CGCGCTCGAG-TTTGAAAATATAGGTATCAGG Porward AAAGAATTC-GACAGAATCGGCGAACGCCGCCCCGC Porward AAAGAATTC-GACAGAATCGGCGAACGCCTCAACCCC Porward AAAGAATTC-GACAGAATCGGCGAACGCCCCCCCCCCCCC	007	Reverse CCCGCTCGAG-CCACTGATAACCGACAGAT	XhoI
Forward AAAGAATTC-GCAGAGTTTGCGCGGGGAAACTTAT  Reverse AAACTGCAG-TTAATATGGTTTTGTCGTTCG  Pst I  Pop Forward CGCGGATCCCATATG-TGCGCGTGGAAACTTAT  Reverse CCCGCTCGAG-TCGGTTTTGAAACTTTGGTTTT  Reverse CCCGCTCGAG-TCGGTTTTGAAACTTTGGTTTT  Pop Forward AAAGAATTC-GCATTTGCCGCGCGAGCG  Reverse AAACTGCAG-TCAGCGATCGAGCTGCTCTTT  Pst I  Forward AAAGAATTC-GCTTTCCGCGTGGCCGGCGTGC  Reverse AAAAAACTGCAG-GTCGACTTATTCGGCGGCTTTTTCCGC  Pst I  Forward AAAAAAAAACTC-CCACTCAAGTATTGAG  Reverse AAAAAACTGCAG-TTACAGTCCGTCCACGCCTTTCCGC  Pst I  Forward CGCGGATCCCATATG-GAAACCCCCCCCGC  BamHI-  Reverse CCCGCTCGAG-AGGTTGTTTCCAGGTTG  Phy I  Reverse CCCGCTCGAG-AGGTTGTTCCAGGTTG  Pop I  Reverse CCCGCTCGAG-AGGTTGTTCCAGGTTG  Pop I  Reverse CCCGCTCGAG-TTTGAAAACGCCAGCGGAA  Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG  Pop I  Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG  Pop I  Reverse CCCGCTCGAG-TTTGAAAAATATAGGTATCAGG  Pop I  Reverse CCCGCTCGAG-TTTGAAAATTATGGTATCAGG  Pop I  Reverse CCCGCTCGAG-TTTGAAAAATATAGGTATCAGG  Pop I  Reverse CCCGCTCGAG-TTTTGAAAAATATAGGTATCAGG  Pop I  Reverse CCCGCTCGAG-TTTTGAAAATTATGGTATCAGG  Pop I  Reverse AAACTGCAG-CTATATGCGCGGCAAGGACGCGCACGCACGACGCACGACGACGACG	907	Forward CGCGGATCCCATATG-GGCGCGCAACGTGAG	BamHI-
Forward AAAGAATTC-GCAGAGTTAGTAGGCGTTAATAAAAATAC Reverse AAACTGCAG-TTAATATGGTTTTGTCGTTCG Pst I  909 Forward CGCGGATCCCATATG-TGCGCGTGGGAAACTTAT  Reverse CCCGCTCGAG-TCGGTTTTGAAACTTTGGTTTT Reverse CCCGCTCGAG-TCGGTTTTGAAACTTTGGTTTT  910 Forward AAAGAATTC-GCATTTGCCGGCGAGCG Reverse AAACTGCAG-TCAGCGATCGAGCTGCTCTTT  911 Forward AAAGAATTC-GCTTTCCGCGTGGCCGGCGGTGC Reverse AAAAAACTGCAG-GTCGACTTATTCGGCGGCGTTTTCCGC Pst I  912 Forward AAAAAAGAATTC- CAAATCCGTCAAAACGCCACTCAAGTATTGAG Reverse AAAAAACTGCAG-TTACAGTCCGTCCACGCCTTTCGC Pst I  913 Forward CGCGGATCCCATATG-GAAACCCGCCCCGC BamHI- Ndel Reverse CCCGCTCGAG-AGGTTGTTTCCAGGTTG Ndel Shol Pit Forward AAAGAATTC-GCCGGCAGGCGGAA Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG Pit Forward AAAGAATTC-GACAGAATCGGCGATTTGGAAGCACG Reverse AAACTGCAG-CTATATGCGCGGCAAGGACGCC Reverse AAACTGCAG-CTATATGCGCGGCAAGGACCGCCCCGC Reverse AAACTGCAG-CTATATGCGCGGCAAGGACGCCCCCCGC Reverse AAACTGCAG-CTATATGCGCGGCAAGCACGCCCCCCCCCC		Payrama CCCCCTCCAC ACCCCACACACACACACACACACACACA	NdeI
Reverse AAACTGCAG-TTAATATGGTTTTGTCGTTCG  909 Forward CGCGGATCCCATATG-TGCGCGTGGGAAACTTAT  Reverse CCCGCTCGAG-TCGGTTTTGAAACTTTGGTTTT  910 Forward AAAGAATTC-GCATTTGCCGGCGAGCG  Reverse AAACTGCAG-TCAGCGATCGAGCTGCTCTTT  911 Forward AAAGAATTC-GCTTTCCGCGTGGCCGGCGTGC  Reverse AAAAAACTGCAG-GTCGACTTATTCGGCGGCGTTTTCCGC  912 Forward AAAAAAACTGCAG-GTCGACTTATTCGGCGGCTTTTCCGC  Reverse AAAAAAACTGCAG-TTACAGTCCGTCCACGCCTTTCGC  913 Forward CGCGGATCCCATATG-GAAACCCGCCCCGC  914 Forward AAAGAATTC-  Reverse CCCGCTCGAG-AGGTTGTTCCAGGTTG  915 Forward CGCGGATCCCATATG-TGCCGGCAGGCGGAA  8 BamHI-  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  NdeI  Shoi  Forward CGCGGATCCCATATG-TGCCGGCAGGCGGAA  BamHI-  NdeI  Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG  914 Forward AAAGAATTC-GACAGAATCGGCGATTTGGAAGCACG  Reverse AAACTGCAG-CTATATGCGCGGCAGGACGCTCAACGC  Reverse AAACTGCAG-CTATATGCGCGGCAGGACGCTCAACGCC  Reverse AAACTGCAG-CTATATGCGCGGCAGGCGCTCAACGCC  Reverse AAACTGCAG-CTATATGCGCGGCAGGCGCTCAACGCC  Reverse AAACTGCAG-CTATATGCGCGGCAGGCGCTCAACGCC  Reverse AAACTGCAG-CTATATGCGCGGCAGGCGCTCAACGCCC  Reverse AAACTGCAG-CTATATGCGCGGCAGGCGCTCAACGCC  Reverse AAACTGCAG-CTATATGCGCGCGCAGGCGCTCAACGCCCCCCCCCC	000	Reverse CCCGCTCGAG-ACGCCACTGCCAGCG	XhoI
Porward CGCGGATCCCATATG-TGCGCGTGGGAAACTTAT  Reverse CCCGCTCGAG-TCGGTTTTGAAACTTTGGTTTT  Poward AAAGAATTC-GCATTTGCCGGCGAGCG  Reverse AAACTGCAG-TCAGCGATCGAGCTGCTCTTT  Pst I  Poward AAAGAATTC-GCTTTCCGCGTGGCGGCGGTGC  Reverse AAAAAACTGCAG-GTCGACTTATTCGGCGGCGTGC  Reverse AAAAAACTGCAG-GTCGACTTATTCGGCGGCGTTTTTCCGC  Pst I  CAAATCCGTCAAAACGCCACTCAAGTATTGAG  Reverse AAAAAACTGCAG-TTACAGTCCGTCCACGCCTTCCGC  Pst I  Forward CGCGGATCCCATATG-GAAACCCGCCCCGC  BamHI-  Ndel  Reverse CCCGCTCGAG-AGGTTGTTCCAGGTTG  Proward CGCGGATCCCATATG-TGCCGGCAGGCGAA  Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG  Poward AAAGAATTC-GACAGAATCGGCGATTTGGAAGCACG  Reverse AAACTGCAG-CTATATGCGCGGCAGGACGCTCAACGC  Reverse AAACTGCAG-CTATATGCGCGGCAGGACGCTCAACGC  Reverse AAACTGCAG-CTATATGCGCGGCAGGACGCTCAACGC  Reverse AAACTGCAG-CTATATGCGCGGCAGGACGCTCAACGC  Reverse AAACTGCAG-CTATATGCGCGGCAGGACGCTCAACGC  Reverse AAACTGCAG-CTATATGCGCGGCAGGACGCTCAACGC  Reverse AAACTGCAG-CTATATGCGCGGCAGGACGCTCAACGC  Reverse AAACTGCAG-CTATATGCGCGGCAGGACGCTCAACGCC  Reverse AAACTGCAG-CTATATGCGCGGCAGGACGCTCAACGCC  Reverse AAACTGCAG-CTATATGCGCGGCAGGACGCTCCAACGCC  Reverse AAACTGCAG-CTATATGCGCGGCAGGACGCTCCAACGCC  Reverse AAACTGCAG-CTATATGCGCGGCAGGACGCTCCAACGCC  Reverse CCCGCTCGAG-CTATATGCGCGCGCAGGACGCCTCCAACGCC  Reverse AAACTGCAG-CTATATGCGCGCGCAGGACGCCTCCAACGCC  Reverse CCCGCTCGAG-CTATATGCGCGCGCAGGACGCCTCCAACGCC  Reverse AAACTGCAG-CTATATGCGCGCGCAGGACGCCTCCAACGCC  Reverse CCCGCTCGAG-CTATATGCGCGCGCAGGACGCCTCCAACGCC  Reverse AAACTGCAG-CTATATGCGCGCGCAGGACGCTCCAACGCC  Reverse CCCGCTCGAG-CTATATGCGCCGCCTCCAACGCC  Reverse CCCGCTCGAG-CTCTCAACGCC  Reverse CCCGCTCGAG-CTCCAACGCC  Reverse CCCGCTCGAG-CTCCAACGCC  Reverse CCCGCTCGAG-CTCCAACGCC  Reverse CCCGCTCGAG-CTCCAACGCC  Reverse CCCGCTCGAG-CTCCAACGCC  Reverse AAACTGCAG-CTCTCAACGCC  Reverse CCCGCTCGAG-CTCCAACGCC  Reverse CCCGCTCCACGCCTCCAACGCC  Reverse CCCGCTCCACGCCTCCACCCCCCCCCCCCCCCCCCCCCC	900	Poward AAAGAATIC-GCAGAGTTAGTAGGCGTTAATAAAAATAC	Eco RI
Reverse CCCGCTCGAG-TCGGTTTTGAAACTTTGGTTTT  910 Forward AAAGAATTC-GCATTTGCCGGCGACTCTGCCGAGCG Reverse AAACTGCAG-TCAGCGATCGAGCTGCTCTTT  911 Forward AAAGAATTC-GCTTTCCGCGTGGCCGGCGGTGC Reverse AAAAAACTGCAG-GTCGACTTATTCGGCGGCGTTTTTCCGC Pst I  912 Forward AAAAAAACTGCAG-GTCGACTTATTCGGCGGCTTTTTCCGC Reverse AAAAAAACTGCAG-TTACAGTCCGTCCACGCCTTTCGC Reverse AAAAAAACTGCAG-TTACAGTCCGTCCACGCCTTTCGC Pst I  913 Forward CGCGGATCCCATATG-GAAACCCGCCCCGC BamHI- Ndel Reverse CCCGCTCGAG-AGGTTGTTCCAGGTTG XhoI  Proward CGCGGATCCCATATG-TGCCGGCAGGCGGAA  Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG Severse AAACTGCAG-TTTGAAAATTTGGGCGGAAGCACG Reverse AAACTGCAG-CTATATGCGCGGCAGGCGCTCAACGC  Political Notes Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Short Sho	000	Reverse AAACIGCAG-TTAATATGGTTTTGTCGTTCG	Pst I
Reverse CCCGCTCGAG-TCGGTTTTGAAACTTTGGTTTT  910 Forward AAAGAATTC-GCATTTGCCGGCGACTCTGCCGAGCG Reverse AAACTGCAG-TCAGCGATCGAGCTGCTCTTT  911 Forward AAAGAATTC-GCTTTCCGCGTGGCCGGCGGTGC Reverse AAAAAACTGCAG-GTCGACTTATTCGGCGGCGTTTTTCCGC  912 Forward AAAAAACTGCAG-GTCGACTTATTCGGCGGCTTTTTCCGC Pst I  CAAATCCGTCAAAAACGCCACTCAAGTATTGAG Reverse AAAAAACTGCAG-TTACAGTCCGTCCACGCCTTTCGC Pst I  913 Forward CGCGGATCCCATATG-GAAACCCGCCCCGC  Reverse CCCGCTCGAG-AGGTTGTTTCCAGGTTG  915 Forward CGCGGATCCCATATG-TGCCGGCAGGCGGAA  Reverse CCCGCTCGAG-TTTGAAAAATATAGGTATCAGG Poward AAAGAATTC-GACAGAATCGGCGATTTGGAAGCACG Reverse AAACTGCAG-CTATATGCGCGGCAGGACGCTCAACGC  Pcc Ri  Shol  Shol  Reverse CCCGCTCGAG-TTTGAAAATTATGGTATCAGG Shol  Reverse AAACTGCAG-CTATATGCGCGGCAGGACGCTCAACGC Reverse AAACTGCAG-CTATATGCGCGGCAGGACGCTCAACGC  Pcc Ri  Shol  Sco Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC Ri  CCC	909	Forward CGCGGATCCCATATG-TGCGCGTGGGAAACTTAT	BamHI-
910 Forward AAAGAATTC-GCATTTGCCGGCGACTCTGCCGAGCG Reverse AAACTGCAG-TCAGCGATCGAGCTGCTCTTT  911 Forward AAAGAATTC-GCTTTCCGCGTGGCCGGCGGTGC Reverse AAAAAACTGCAG-GTCGACTTATTCGGCGGCGTGC Pst I  912 Forward AAAAAAAGAATTC- CAAATCCGTCAAAAACGCCACTCAAGTATTGAG Reverse AAAAAAACTGCAG-TTACAGTCCGTCCACGCCTTTCGC 913 Forward CGCGGATCCCATATG-GAAACCCGCCCCGC Reverse CCCGCTCGAG-AGGTTGTTCCAGGTTG 915 Forward CGCGGATCCCATATG-TGCCGGCAGGCGGAA  Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG 916 Forward AAAGAATTC-GACAGAATCGGCGATTTGGAAGCACG Reverse AAACTGCAG-CTATATGCGCGCGCAGGCAGACGCCCCCGC Reverse AAACTGCAG-CTATATGCGCGCGCAGGCACGCCTCAACGC  Pot I  Eco RI  Ndel  Ndel  Ndel  Ndel  Ndel  Reverse CCCGCTCGAG-TTTGAAAAATATAGGTATCAGG Phil		Peverse CCCCCTCCAC TGGGTTTTTG	NdeI
Reverse AAACTGCAG-TCAGCGATCGAGCTGCTCTTT  911 Forward AAAGAATTC-GCTTTCCGCGTGGCCGGCGGTGC  Reverse AAAAAACTGCAG-GTCGACTTATTCGGCGGCGGTTTTTCCGC  912 Forward AAAAAAAGAATTC-  CAAATCCGTCAAAAACGCCACTCAAGTATTGAG  Reverse AAAAAAACTGCAG-TTACAGTCCGTCCACGCCTTTCGC  913 Forward CGCGGATCCCATATG-GAAACCCGCCCCGC  Reverse CCCGCTCGAG-AGGTTGTTCCAGGTTG  915 Forward CGCGGATCCCATATG-TGCCGGCAGGCGGAA  Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG  916 Forward AAAGAATTC-GACAGAATCGGCGATTTGGAAGCACG  82 Reverse AAACTGCAG-CTATATGCGCGGCACGCACGCCCCCGC  Reverse AAACTGCAG-CTATATGCGCGGCACGCACGCCCCCAACCGC  8 Eco RI  8 CO RI	010	Forward AAAGAATTC CCATTTTGGTTTT	XhoI
911 Forward AAAGAATTC-GCTTTCCGCGTGGCCGGCGGTGC Reverse AAAAAACTGCAG-GTCGACTTATTCGGCGGCTTTTTCCGC Pst I CAAATCCGTCAAAACGCCACTCAAGTATTGAG Reverse AAAAAACTGCAG-TTACAGTCCGTCCACGCCTTTCGC Pst I Forward CGCGGATCCCATATG-GAAACCCGCCCCGC BamHI- Ndel Reverse CCCGCTCGAG-AGGTTGTTCCAGGTTG Sorward CGCGGATCCCATATG-TGCCGGCAGGCGGAA Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG Pst I Strict Pst I Sco Ri Ndel Ndel Ndel Ndel Ndel Reverse CCCGCTCGAG-AGGTTGTTCCAGGTTG Sorward CGCGGATCCCATATG-TGCCGGCAGGCGGAA Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG Pst I Sco Ri Sco Ri Sco Ri Sco Ri Sco Ri Sco Ri Sco Ri Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG Sorward AAAGAATTC-GACAGAATCGGCGATTTGGAAGCACG Reverse AAACTGCAG-CTATATGCGCGGCAGGACGCTCAACGC	710	Powers AAAGAATIC-GCATTIGCCGGCGACTCTGCCGAGCG	Eco RI
Reverse AAAAAACTGCAG-GTCGACTTATTCGGCGGCTTTTTCCGC  912 Forward AAAAAAGAATTC- CAAATCCGTCAAAAACGCCACTCAAGTATTGAG Reverse AAAAAAACTGCAG-TTACAGTCCGTCCACGCCTTTCGC  913 Forward CGCGGATCCCATATG-GAAACCCGCCCCGC  Reverse CCCGCTCGAG-AGGTTGTTCCAGGTTG  915 Forward CGCGGATCCCATATG-TGCCGGCAGGCGGAA  Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG  916 Forward AAAGAATTC-GACAGAATCGGCGATCTCAACGG  917 Forward AAAGAATTC-GACAGAATCGGCGAACGACGCCTCAACGG  918 I Eco RI  Ndel Xhol BamHI- Ndel Xhol Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST I Eco RI  PST	011	Reverse AAACIGCAG-ICAGCGATCGAGCTGCTCTTT	Pst I
CAAATCCGTCAAAACGCCACTCAAGTATTGAG Reverse AAAAAACTGCAG-TTACAGTCCGTCCACGCCTTTCGC  913 Forward CGCGGATCCCATATG-GAAACCCGCCCCGC Reverse CCCGCTCGAG-AGGTTGTTCCAGGTTG  915 Forward CGCGGATCCCATATG-TGCCGGCAGGCGGAA  Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG  916 Forward AAAGAATTC-GACAGAATCGGCGATTGGAAGCACG  917 Forward AAAGAATTC-GACAGAATCGGCGAACGCACGCCCCCCCCCC	911	Forward AAAGAATTC-GCTTTCCGCGTGGCCGGCGGTGC	Eco RI
CAAATCCGTCAAAACGCCACTCAAGTATTGAG  Reverse AAAAAACTGCAG-TTACAGTCCGTCCACGCCTTTCGC  913 Forward CGCGGATCCCATATG-GAAACCCGCCCCGC  Reverse CCCGCTCGAG-AGGTTGTTCCAGGTTG  915 Forward CGCGGATCCCATATG-TGCCGGCAGGCGGAA  Reverse CCCGCTCGAG-TTTGAAAAATATAGGTATCAGG  916 Forward AAAGAATTC-GACAGAATCGGCGATTTGGAAGCACG  Reverse AAACTGCAG-CTATATGCGCGGCAGGACGCTCAACGC  Eco RI  Eco RI  Pst I  BamHI- NdeI  XhoI  Eco RI  Eco RI  Eco RI	010	Reverse AAAAAACTGCAG-GTCGACTTATTCGGCGGCTTTTTCCGC	Pst I
Reverse AAAAAACTGCAG-TTACAGTCCGTCCACGCCTTTCGC  913 Forward CGCGGATCCCATATG-GAAACCCGCCCCGC  Reverse CCCGCTCGAG-AGGTTGTTCCAGGTTG  915 Forward CGCGGATCCCATATG-TGCCGGCAGGCGGAA  Reverse CCCGCTCGAG-TTTGAAAAATATAGGTATCAGG  916 Forward AAAGAATTC-GACAGAATCGGCGATTTGGAAGCACG  Reverse AAACTGCAG-CTATATGCGCGGCAGGACGCTCAACGC  Pst I  NdeI  NdeI  XhoI  Eco RI  Reverse AAACTGCAG-CTATATGCGCGGCAGGACGCTCAACGC	912	Forward AAAAAAGAATTC-	
Porward CGCGGATCCCATATG-GAAACCCGCCCCGC  Reverse CCCGCTCGAG-AGGTTGTTCCAGGTTG  915 Forward CGCGGATCCCATATG-TGCCGGCAGGCGGAA  Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG  914 Forward AAAGAATTC-GACAGAATCGGCGATTTGGAAGCACG  Reverse AAACTGCAG-CTATATGCGCGGCAGGACGCTCAACGC  BamHI- Ndel Xhol Eco RI Eco RI		Payore AAAAAACTCAACTATTGAG	
Reverse CCCGCTCGAG-AGGTTGTTCCAGGTTG  915 Forward CGCGGATCCCATATG-TGCCGGCAGGCGGAA  Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG  914 Forward AAAGAATTC-GACAGAATCGGCGATTTGGAAGCACG  Reverse AAACTGCAG-CTATATGCGCGGCAGGACGCTCAACGC  Eco RI	012	Forward CCCCCATGGGATTACAGTCCGTCCACGCCTTTCGC	Pst I
Provided CCCGCTCGAG-AGGTTGTTCCAGGTTG  915 Forward CGCGGATCCCATATG-TGCCGGCAGGCGGAA  Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG  914 Forward AAAGAATTC-GACAGAATCGGCGATTTGGAAGCACG  Reverse AAACTGCAG-CTATATGCGCGGCAGGACGCTCAACGC  Eco RI	913	Forward CGCGGATCCCATATG-GAAACCCGCCCCGC	BamHI-
915 Forward CGCGGATCCCATATG-TGCCGGCAGGCGGAA  Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG  914 Forward AAAGAATTC-GACAGAATCGGCGATTTGGAAGCACG  Reverse AAACTGCAG-CTATATGCGCGCGCAGGACGCTCAACGC  Eco RI		Reverse CCCCCTCCAC ACCTTCTCTCTCTC	NdeI
Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG  NdeI  ShoI  Forward AAAGAATTC-GACAGAATCGGCGATTTGGAAGCACG  Reverse AAACTGCAG-CTATATGCGCGGCAGGACGCTCAACGC  Reverse AAACTGCAG-CTATATGCGCGCGCAGGACGCTCAACGC	015	Forward CCCCCATCCCATCCCATCATCATCATCATCATCATCATCA	XhoI
Proverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG  Sholl Forward AAAGAATTC-GACAGAATCGGCGATTTGGAAGCACG  Reverse AAACTGCAG-CTATATGCGCGCAGGACGCTCAACGC  Reverse AAACTGCAG-CTATATGCGCGCGCAGGACGCTCAACGC	913	Forward CGCGGATCCCATATG-TGCCGGCAGGCGGAA	BamHI-
Porward AAAGAATTC-GACAGAATCGGCGATTTGGAAGCACG  Reverse AAACTGCAG-CTATATGCGCGCAGGACGCTCAACGC  Eco RI		Reverse CCCGCTCGAG TTTCAAAAMATA	NdeI
Reverse AAACTGCAG-CTATATGCGCGCAGGACGCTCAACGC	914	Forward AAAGAATTC CACACAATTAGGTATCAGG	XhoI
916 Forward CGCGGATCGCATATG GGALATTE	7 <del>1 T</del>	Reverse AAACTGCAC CTATATICS CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACTOR AND CONTRACT	Eco RI
	916	Forward CCCCCATCCCATATO	Pst I
916 Forward CGCGGATCCCATATG-GCAATGATGGCGGCTG  Pst 1  BamHI-	710	TOIWAIL COCOGNICCCATATG-GCAATGATGGCGGCTG	BamHI-
374 -		Reverse CCCGCTCGAG TTTCCCCCGG TTTTCCCCCGG	
Reverse CCCGCTCGAG-TTTGGCGGCATCTTTCAT Ndel XhoI		TOTAL COCOCTOMO-111GGCGGCATCTTCAT	XhoI

917	E	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	
91/		AAAAAAGAATTC-CCTGCCGAAAAACCGGCACCGGC	Eco RI
010		AAAAAACTGCAG-TTATTTCCCCGCCTTCACATCCTG	Pst I
919	Forward	CGCGGATCCCATATG-TGCCAAAGCAAGAGCATC	BamHI-
	Darrama	CCCCCTCC A C CCCCCCCT A TTTTCCCC	NdeI
920		CCCGCTCGAG-CGGGCGGTATTCGGG	XhoI
920	Forward	CGCGGATCCCATATG-CACCGCGTCTGGGTC	BamHI-
	Dovorco	CCCGCTCGAG-ATGGTGCGAATGACCGA	NdeI
021			XhoI
921		AAAAAAGAATTC-TTGACGGAAATCCCCGTGAATCC	Eco RI
022	Reverse	AAAAAACTGCAG-TCATTTCAAGGGCTGCATCTTCAT	Pst I
922		CGCGGATCCGCTAGC-TGTACGGCGATGGAGGC	BamHI-
	2 Payarsa	CCCGCTCGAG-CAATCCCGGGCCGCC	NheI
923			XhoI
723	roiwaiu	CGCGGATCCCATATG-TGTTACGCAATATTGTCCC	BamHI-
	Reverse	CCCGCTCGAG-GGACAAGGCGACGAAG	NheI
925		CGCGGATCCCATATG-AAACAAATGCTTTTAGCCG	XhoI
745	1 OI Wald	COCOGNICCCATATO-AAACAAATGCTTTTAGCCG	BamHI-
	Reverse	CCCGCTCGAG-GCCGTTGCATTTGATTTC	Ndel
926		CGCGGATCCCATATG-TGCGCGCAATTACCTC	XhoI
	1 OI Ward	-IdededeAlliAcele	BamHI-
	Reverse	CCCGCTCGAG-TCTCGTGCGCGCCG	NdeI XhoI
927		CGCGGATCCCATATG-TGCAGCCCCGCAGC	BamHI-
			Ndel
	Reverse	CCCGCTCGAG-GTTTTTTGCTGACGTAGT	XhoI
929a	Forward	AAAAAAGAATTC-CGCGGTTTGCTCAAAACAGGGCTGGG	Eco RI
	Reverse	AAAAAATCTAGA-TTAAGAAAGACGGAAACTACTGCC	Xba I
931	Forward	AAAAAAGAATTC-GCAACCCATGTTTTGATGGAAAC	Eco RI
	Reverse	AAAAAACTGCAG-TTACTGCCCGACAACAACGCGACG	Pst I
935	Forward	AAAAAAGAATTC-	Eco RI
		GCGGATGCGCCCGCGATTTTGGATGACAAGGC	Loo Id
	Reverse	AAAAAACTGCAG-TCAAAACCGCCAATCCGCCGACAC	Pst I
936	Forward	CGCGGATCCCATATG-GCCGCCGTCGGCGC	BamHI-
	_		NdeI
		CCCGCTCGAG-GCGTTGGACGTAGTTTTG	XhoI
937	Forward .	AAAAAAGAATTC-CCGGTTTACATTCAAACCGGCGCAAC	Eco RI
	Reverse	AAAAAA <u>CTGCAG</u> -TTAAAATGTATGCTGTACGCCAAA	Pst I
939a		AAAAAAGAATTC-GGTTCGGCAGCTGTGATGAAACC	Eco RI
	Reverse	AAAAAA <u>CTGCAG</u> -TTAACGCAAACCTTGGATAAAGTTGGC	Pst I
950	Forward	CGCGGATCCCATATG-GCCAACAACCGGCAAG	BamHI-
	D	GGGGGTTGG + G ==== + G	NdeI
0.50		CCCGCTCGAG-TTTAGAACCGCATTTGCC	XhoI
953	Forward	CGC <u>GGATCCCATATG</u> -GCCACCTACAAAGTGGAC	BamHI-
	Davors		NdeI
057		CCCGCTCGAG-TTGTTTGGCTGCCTCGAT	XhoI
957	rorward	CGC <u>GGATCCCATATG</u> -TTTTGGCTGGGAACGGG	BamHI-

		3717
	Reverse CCCGCTCGAG-GTGAGAAAGGTCGCGC	Ndel
958	Forward CGCGGATCCCATATG-GCCGATGCCGTTGCG	XhoI
"	Toward CocooniceCatato. GCCGATGCCGTTGCG	BamHI-
	Reverse GCCCAAGCTT-GGGTCGTTTGTTGCGTC	NdeI
959	Forward CGCGGATCCCATATG-CACCACGACGGACACG	HindIII
	Torward COCOGATECCATATO-CACCACGACGCGACACG	BamHI-
	Reverse CCCGCTCGAG-GTCGTCGCGGGGGA	NdeI
961	Forward CGCGGATCCCATATG-GCCACAAGCGACGACG	XhoI
	TOWARD COCONTAIN-OCCACAAGCGACGACG	BamHI-
	Reverse CCCGCTCGAG-CCACTCGTAATTGACGC	NdeI
972	Forward AAAAAAGAATTC-	XhoI
	TTGACTAACAGGGGGGGGGGGGAAATTAAAAAC	Eco RI
	Reverse AAAAAATCTAGA-TTAAAAATAATCATAATCTACATTTTG	<b>371</b> •
973	Forward AAAAAAGAATTC-ATGGACGCGCACAACCGAAAAC	Xba I
	Reverse AAAAAACTGCAG-TTACTTCACGCGGGTCGCCATCAGCGT	Eco RI
982	Forward CGCGGATCCCATATG-GCAGCAAAAGACGTAC	Pst I
	- GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD TO GOOD	BamHI-
	Reverse CCCGCTCGAG-CATCATGCCGCCCATCC	NdeI
983	Forward CGCGGATCCCATATG-TTAGCTGTTGCAACAACAC	XhoI
		BamHI-
	Reverse CCCGCTCGAG-GAACCGGTAGCCTACG	NdeI XhoI
987	Forward CGCGGATCCCATATG-CCCCCACTGGAAGAAC	BamHI-
		NdeI
•	Reverse CCCGCTCGAG-TAATAAACCTTCTATGGGC	XhoI
988	Forward CGCGGATCCCATATG-TCTTTAAATTTACGGGAAAAAG	BamHI-
	_	NdeI
	Reverse GCCCAAGCTT-TGATTTGCCTTTTCCGTTTT	HindIII
989 1	Forward CCGGAATTCTACATATG-GTCCACGCATCCGGCTA	EcoRI-
		NdeI
000 5	Reverse CCCGCTCGAG-TTTGAATTTGTAGGTGTATTGC	XhoI
990 F	Forward. CGCGGATCCGCTAGC-TTCAGAGCTCAGCTT	BamHI-
1	Persona CCCCCTCCAC AAAGA GGG	NheI
002 1	Reverse CCCGCTCGAG-AAACAGCCATTTGAGCGA	XhoI
<i>974</i> [	Forward CGCGGATCCCATATG-GACGCGCCCGCCCG	BamHI-
ı	Reverse CCCGCTCGAG-CCAAATGCCCAACCATTC	NdeI
993 F	Forward CGCGGATCCCATATG-GCAATGCTGATTGAAATCA	XhoI
,,,,	OWAR COCOGNICCCATATO-OCAATGCTGATTGAAATCA	BamHI-
F	Reverse CCCGCTCGAG-GAACACATCGCGCCCG	Ndel
996 F	Forward CGCGGATCCCATATG-TGCGGCAGAAAATCCGC	XhoI
	- I GCGGCAGAAAICCGC	BamHI-
F	Reverse CCCGCTCGAG-TCTAAACCCCTGTTTTCTC	NdeI
997 F	Forward CCGGAATTCTAGCTAGC-CGGCACGCCGACGTT	XhoI
	COULTED COOCACOCCOACGII	EcoRI-
F	Reverse CCCGCTCGAG-GACGGCATCGCTCAGG	NheI
		XhoI

4.5

The second second second second second second second second second second second second second second second se

Underlined sequences indicate restriction recognition sites.

```
The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 1>:
     g001.seq
              ATGCTGCCGC AGGGGAAGGC GGCGCGGAGG GTGTCGGCGA ACGAGGTGTC
           1
          51 CGGCAGGCT TGCGCCCGGA TGGTGCTGGT CATCTGCCAG ACGCTGCCGA
          101 AACGCGATAC TTTAAACGGC TCGGGTACGC ATACTTTACC GGTTTGGGCG
          151 ATTTTGCCGA GGTCGTTGCG CAGCAAATCG ACAATCATCA CGTTTTCGGC
          201 GCGGTTTTTC GGGTCGGTTT GTAACTCGGC GGCGCGGCGT TCGTCTTGTC
              301
              CCGTCTGAAG CGATGTTGAG GAAGAGTTCG GGCGAGAAAC ACAGCGTCCA
         351 CGCGGATTGC CCGGCTTCAT CGGGCAGGTG GGACAATACG GCATAG
This corresponds to the amino acid sequence <SEQ ID 2; ORF 001.ng>:
     g001.pep
              MLPQGKAARR VSANEVSGRA CARMVLVICO TLPKRDTLNG SGTHTLPVWA
              ILPRSLRSKS TIITFSARFF GSVCNSAARR SSCPSPKIGA VPFIGSVLMV
         101 PSEAMLRKSS GEKHSVHADC PASSGRWDNT A*
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 3>:
     m001.seq
             ATGCTGCCGC AGGGGAAGGC GGCGCGGAGG ATGTCGGCGA ACGAGGTGTG
              CGGCAssCTT ss.GCTTGGA YGGTGCTGGT CATCTGCCAA ACGCTGCCGA
         101 AACGCGATAC TTTAAACGGT TCGGGTACGC ATACTGTGCC GGTTTGGGCG
         151 ATTTTGCCGA GATCGTTACG CAGCAAATCG ACAATCATCA CGTTTTCGGC
         201 GCGGTTTTTC GGGTCTGCTT GCAACTCGGC GGCGCGGCGT TCGTCTTGTC
         CCGTCCGAAC CGATTTTGAG GAAGAGTTCG GGCGAGAAAC ACAGCGTCCA
              CGCGGATTGC CCCTCCGCAT CGGGCAGGTG GGACAAGACG GCATAG
This corresponds to the amino acid sequence <SEQ ID 4; ORF 001>:
    m001.pep
              MLPQGKAARR MSANEVCGXL XAWXVLVICQ TLPKRDTLNG SGTHTVPVWA
          51
              ILPRSLRSKS TIITFSARFF GSACNSAARR SSCPSPKIGA VPFIGSVLMV
         101 PSEPILRKSS GEKHSVHADC PSASGRWDKT A*
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 5>:
     a001.seq
              ATGCTGCCGC AGGGGAAGGC GGCGCGGAGG ATGTCGGCGA ACGAGGTGTG
              CGGCAAGGCT TGGGCTTGGA TGGTGCTGGT CATCTGCCAA ACGCTGCCGA
              AACGCGATAC TTTAAACGGT TCGGGTACGC ATACTGTGCC GGTTTGGGCG
         151 ATTTTGCCGA GGTCGTTACG CAGCAAATCG ACAATCATCA CGTTTTCGGC
         201 GCGGTTTTTC GGGTCTGCTT GCAACTCGGC GGCGCGGCGT TCGTCTTGTC
         CCGTCCGAAC CGATTTGAG GAAGAGTTCG GGCGAGAAAC ACAGCGTCCA
         351 CGCGGATTGC CCTTGTGCAT CGGGCAGGTG GGACAAAACG GCATAG
This corresponds to the amino acid sequence <SEQ ID 6; ORF 001.a>:
    a001.pep
              MLPQGKAARR MSANEVCGKA WAWMVLVICQ TLPKRDTLNG SGTHTVPVWA
              ILPRSLRSKS TIITFSARFF GSACNSAARR SSCPSPKIGA VPFIGSVLMV
         101 PSEPILRKSS GEKHSVHADC PCASGRWDKT A*
    m001/a001
                96.2% identity over a 131 aa overlap
                                 20
                                                    40
                                                             50
                MLPQGKAARRMSANEVCGXLXAWXVLVICQTLPKRDTLNGSGTHTVPVWAILPRSLRSKS
    m001.pep
                                   1111111111111111
                MLPQGKAARRMSANEVCGKAWAWMVLVICQTLPKRDTLNGSGTHTVPVWAILPRSLRSKS
    a001.pep
```

BNSDOCID: <WO___9957280A2_j_>

	10	20	30	40	50	60
m001.pep	70 TIITFSARFFGSACN           TIITFSARFFGSACN 70			111111111	1111111111	11111
m001.pep	130 PSASGRWDKTAX            PCASGRWDKTAX 130					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae

ORF 001 shows 89.3% identity over a 131 aa overlap with a predicted ORF (ORF 001.ng) from N. gonorrhoeae:

m001/g001

```
20
                                30
                                        40
                                                50
          MLPQGKAARRMSANEVCGXLXAWXVLVICQTLPKRDTLNGSGTHTVPVWAILPRSLRSKS
m001.pep
          MLPQGKAARRVSANEVSGRACARMVLVICQTLPKRDTLNGSGTHTLPVWAILPRSLRSKS
g001
                10
                        20
                                30
                                                50
                                                        60
                70
                        80
                                90
                                       100
                                               110
          TIITFSARFFGSACNSAARRSSCPSPKIGAVPFIGSVLMVPSEPILRKSSGEKHSVHADC
m001.pep
          g001
          TIITFSARFFGSVCNSAARRSSCPSPKIGAVPFIGSVLMVPSEAMLRKSSGEKHSVHADC
                70
                        80
                                90
                                       100
                130
m001.pep
          PSASGRWDKTAX
          1::11111:111
g001
          PASSGRWDNTAX
               130
```

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 7>:

```
ATGGTCGTAT TCGTGGCTGA AGGCGTATTC GGTCGCGCTG TTTTGGGTCA
    CTTGGTATTG CTCTTCGGTC AGGGTGCGTT TGAGTTCGGC GTCACTCGGT
 51
101
     TTTTTATACG TTGCCGCGTC GAAGCCTTTG CCTTGCGGTG CGGCTTTGGT
151
     TTTGCCCGGC AGCGGTTCGT CGGCTTTGCG GATGTCGATG TGGCAGTAGC
     CGTTGGGGTT TTTAATCAGG TAGTCCTGAT GGTATTCCTC GGCGTCGTAG
201
251 AAGTTTTTCA GCGGTTCGTT TTCAACAACG AGGGGCAGTT GGTATTTTTG
    CTGCTCGCGT TTGAGGGCGG CGGCGATGAC GGCTTTTTCG GCGGGGTCGG
301
351
    TGTAGTACAC GCCGCTGCGG TATTGCGTGC CGGTGTCGTT ACCCTGTTTG
401
    TTGAGGCTGG TCGGATCAAC GACGCGGAAA TAATATTGCA GGATGTCGTC
451
     CAGGCTGagt TTGTCGGCAT CGTaggtcac tTTGACGGTC TCGGCATGAC
     CCGTATGGCG GTaggacact tctTCgtanc TcGGGtTTTC CGTGttGCCG
501
    TTGGCgttac cGGATACCGC gtcaACCACG CCGTcgatgc gttggaAATa
551
     ggCTTCCAAg ccccaaaagc agccgccggc gaagtaaatg gtgcccgtgt
601
651 tcatgattGC TGa
```

This corresponds to the amino acid sequence <SEQ ID 8; ORF 003.ng>: g003.pep

```
1 MVVFVAEGVF GRAVLGHLVL LFGQGAFEFG VTRFFIRCRV EAFALRCGFG
51 FARQRFVGFA DVDVAVAVGV FNQVVLMVFL GVVEVFQRFV FNNEGQLVFL
101 LLAFEGGGDD GFFGGVGVVH AAAVLRAGVV TLFVEAGRIN DAEIILQDVV
```

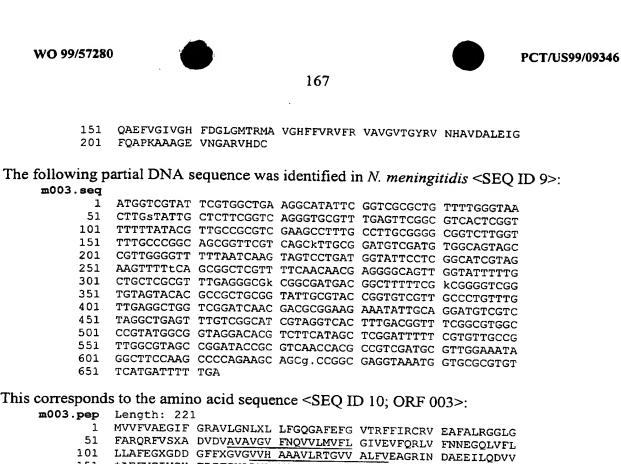
m003.seq

51

301

351

451



### This corresponds to the amino acid sequence <SEQ ID 10; ORF 003>:

```
m003.pep Length: 221
```

- 101 LLAFEGXGDD GFFXGVGVVH AAAVLRTGVV ALFVEAGRIN DAEEILQDVV
- *AEFVGIVGH FDGFGVARMA VGHVFIARIF RVAVGVAGYR VNHAVDALEI
- GFQAPEAAXG EVNGARVHDF *

#### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 11>: a003.seq

```
ATGGTCGTAT TCGTGGCTGA AGGCATATTC GGTCGCGCTG TTTTGGGTAA
 51 CTTGGTATTG CTCTTCGGTC AGGGTGCGTT TGAGTTCGGC GTCACTCGGT
    TTTTTATACG TTGCCGCGTC GAAGCCTTTG CCTTGCGGTG CGGTCTTGGT
    TTTGCCCGGC AGCGGTTCGT CGGCTTTGCG GATATCGATG TGGCAGTAGC
151
201 CGTTGGGGTT TTTAATCAAG TAGTCCTGAT GGTATTCCTC GGCATCGTAG
251 AAGTTTTTCA GCGGCTCGTT TTCAACAACG AGGGGCAGTT GGTATTTTTG
301 CTGCTCGCGT TTGAGGGCGG CGGCGATGAC GGCTTTTTCG GCGGGGTCGG
    TGTAGTACAC GCCGCTGCGG TATTGCGTAC CGGTGTCGTT GCCCTGTTTG
351
    TTGAGGCTGG TCGGATCAAC GACGCGGAAG AAATATTGCA GGATGTCGTC
401
451 TAGGCTGAGT TTGTCGGCAT CGTAGGTCAC TTTGACGGTT TCGGCGTGGC
501 CCGTATGGCG GTAGGACACG TCTTCATAGC TCGGATTTTT CGTGTTGCCG
    TTGGCGTAGC CGGATACCGC GTCAACCACG CCGTCGATGC GTTGGAAATA
551
    GGCTTCCAAG CCCCAGAAGC AGCCGCCGGC GAGGTAGATG GTGCGCGTGT
601
```

TCATGATTTT TGA

#### This corresponds to the amino acid sequence <SEQ ID 12; ORF 003.a>: a003.pep

- MVVFVAEGIF GRAVLGNLVL LFGQGAFEFG VTRFFIRCRV EAFALRCGLG 51 FARQRFVGFA DIDVAVAVGV FNQVVLMVFL GIVEVFQRLV FNNEGQLVFL
- 101 LLAFEGGGDD GFFGGVGVVH AAAVLRTGVV ALFVEAGRIN DAEEILQDVV
- *AEFVGIVGH FDGFGVARMA VGHVFIARIF RVAVGVAGYR VNHAVDALEI
- 201 GFQAPEAAAG EVDGARVHDF *

#### m003/a003 95.9% identity over a 220 aa overlap

	10	20	30	40	50	60
m003.pep	MVVFVAEGIFGRAVI	LGNLXLLFGQ	GAFEFGVTRF	FIRCRVEAFA	LREGUERARO	DEMENA
	11111111111111				DIGGERAQ.	VI ADVV
			1		11 111111	111 - 1
a003	MVVFVAEGIFGRAV]	LGNLVLLFGO	こみ デデ デビリグラ ディ	מים מים עום בים די	I DOGE CENTS	
	• • • • • • • • • • • • • • • • • • • •		OM DI GVINTI	INCKVEALA	PRCCPCLARO	RFVGFA
	10	20	30	4.0	50	60
					50	90

BNSDOCID: <WO 9957280A2 1 >

m003.pep		1   1   1   1   1   7			11 131111	111111
a003	DIDVAVAVGVFNQV	VLMVFLGIV	EVFORLVFNN	EGOLVELLLAF	ECCCDDCFF	
	70	80	90	100	110	120
m003.pep	130	140	150	160	170	180
moos.pep	AAAVLRTGVVALFVI	SAGRINDAE	TLODVVXAEI	VGIVGHFDGF	GVARMAVGH	VFIARIF
a003	AAAVLRTGVVALFVI	::::::::: EAGRINDAEI	IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII		 GVARMAVCHY	
	130	140	150	160	170	180
	190	200	210	220		
m003.pep	RVAVGVAGYRVNHAV	DALEIGFQ <i>i</i>	APEAAXGEVNG	ARVHDFX		
000	111111111111111		1111 111:1	111111		
a003	RVAVGVAGYRVNHA	DALEIGFQ <i>F</i>	PEAAAGEVDG	ARVHDFX		
	190	200	210	220		

Computer analysis of this amino acid sequence gave the following results:

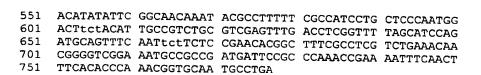
Homology with a predicted ORF from N. gonorrhoeae
ORF 003 shows 88.6% identity over a 219 aa overlap with a predicted ORF (ORF 003.ng) from N. gonorrhoeae:

m003/g003

m003.pep	10	20	30	40	50	60
moos.pep	MVVFVAEGIFGRAVLO	-NLXLLEGQC	AFEFGVTRFF	FIRCRVEAFAL	RGGLGFARQR	FVSXA
g003			111111111	111111111	1 1:11111	11: 1
9003	MVVFVAEGVFGRAVLO	HLVLLFGQG	SAFEFGVTRFF	TIRCRVEAFAL	RCGFGFARQR	FVGFA
	10	20	30	40	50	60
000	70	80	90	100	110	120
m003.pep	DVDVAVAVGVFNQVVI	MVFLGIVEV	FORLVFNNEG	QLVFLLLAFE	GXGDDGFFXG	VGVVH
			111:11111		I IIIII I	11111
g003	DVDVAVAVGVFNQVVI	MVFLGVVEV	FORFVFNNEG	QLVFLLLAFE	GGGDDGFFGG	VGVVH
	<b>7</b> 0	80	90	100	110	120
						120
	130	140	150	160	170	180
m003.pep	AAAVLRTGVVALFVEA	GRINDAEEI	LODVVXAEFV	GIVGHFDGFG	ANDANCHAE.	TABLE
	-	111111111		11111111:1	• • 1 1 1 1 1 1 1 1	: 1:1
g003	AAAVLRAGVVTLFVEA	GRINDAEII	LODVVOAEFV	GIVGHEDGIG	··!!!!   i	:  :
	130	140	150	160	170	
				100	170	180
	190	200	210	220		
m003.pep	RVAVGVAGYRVNHAVD			DUUDEV		
	[][[][][][][][][][][][][][][][][][][][][	111111111	:	NVRDEX		
g003	RVAVGVTGYRVNHAVD	ALETCEOND	·	1111		
J	190	200		RVHDC		
	130	200	210			

# The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 13>:

1	ATGgtagAAC	GGCATATCCA	GCATTTGCGG	AACGGTCATC	TTCATTTCAT
51	GCGCCCATGC	CAACAagtga	gccaAAtgtT	CGGCGGCAGG	GCCTaccatT
101	TCCGCGCCGA	TAAagcggcc	GGTGact TTT	tcgGCataca	occiacyati
151	acctrtatt	ACCAncatca	cacaaataaa	accttgaTTT	ggcgcaratg
201	CTTCCCCcaT	CACAAAmmoo	macaaaaaa	accttgaTTT	TTGAACGATA
251	mammera a a a	GACAAATTCG	TCGGCTTGGT	ATTGCGCGGC	AACCTGCGCG
	TATTTCAAAC	CGACAAAGCC	GATTTGCgga	ctggtaaACA	CCACGCCAAT
301	Geraceded	CGCAAACCGC	TGCCGATAT+	caGataacaa	CCCCCCC++
351	ttgcccggca	atcttacctt	gatcagcagc	ttcatGCAGC	ACCCCCactt
401	ggttggacgc	qtcqcccqca	ataAAGATAT	GCGGAATgct	noodocay:
451	gtCAGCGGAT	CGGCAACGGG	tacaccan	gcgtctttgT	ggtCTGCATg
501	GTTTTCCNNN	CCCAMAM+~m	cacycogege	gcgtctttgT	CGATATTGAT
501	GIIIICCAAA	CCGATATEGT	CAACGTTCGG	ACGGCGACCT	ACGGCTGCCA



## This corresponds to the amino acid sequence <SEQ ID 14; ORF 004.ng>: g004.pep

- 1 MVERHIQHLR NGHLHLMRPC QQVSQMFGGR AYDFRADKAA GGFFGIQAHM
  51 AFVYQHHAAA TLIFERYFAD DKFVGLVLRG NLRVFQTDKA DLRTGKHHAN
  101 GAAAQTAADI RVAAPRYCPA ILPWSAASCS RGSW<u>LDASPA IKICGMLVCM</u>
  151 <u>V</u>SGSATGTPR ASLSILMFSK PILSTFGRRP TAANIYSATN TPFSPSCSQW
  201 TSTLPSASSL TSVLASRCSF NSSPNTAFAS SETTGSEMPP MIPPKPKIST
  251 FTPKRCNA*
- The following partial DNA sequence was identified in N. meningitidis <SEQ ID 15>: m004.seq
  - ATGGTAGAAC GGCATATCCA GCATTTGCGG AACGGTCATC TTCATTTGAT 1 GTGCCCAAGC CAACAGGTGC GCCAAATGTT CGGCGGCAGG GCCTACGATT 51 TCCGCGCCGA TAAAGCGGCC GGTGGCTTTT TCGGCATACA GGCGCATATG 101 151 GCCTTTGTTC ACCAGCATCA CGCGGCTGCG GCCTTGGTTT TTGAACGATA 201 CTTCGCCGAT GACAAATTCG TCGGCTTGGT ATTGCGCGGC AACCTGCGCG 251 TATTTCAGAC CGACAAAGCC GATTTGCGGA CTGGTAAACA CCACGCCGAT 301 GGTGCTGCGC CGCAAACCGC CGCCGATATT CGGGTAGCGG CCGCGTTATC
    351 GCCGGCAATC TTGCCTTGGT CGGCAGCTTC ATGCAGCAGGA GGCAGTTGGT 351 GCCGGCAATC TTGCCTTGGT CGGCAGCTTC ATGCAGCAGA GGCAGTTGGT 401 TGGACGCATC GCCTGCGATG AAGATATGCG GAATACTGGT CTGCATGGTC 451 AGCGGGTCGG CAACAGGTAC GCCGCGCGCA TCTTTTTCGA TATTGATATT 501 TTCCAAACCG ATATTGTCAA CGTTCGGACG GCGGCCCACG GCTGCCAGCA 551 TATATTCGGC AACAAATACG CCTTTTTCGC CATCCTGCTC CCAATGGACT TCTACATTGC CGTCTGCATC GAGTTTGACC TCGGTTTTAG CATCCAGATG 651 CAGTTTCAAT TCTTCGCCGA ACACGGCGTT CGCCTCGTCT GAAACGACGG 701 GGTCGGAAAT GCCGCCGATG ATTCCGCCCA AACCGAAAAT TTCAACTTTC 751 ACGCCCAAAC GGTGCAATGC CTGA

### This corresponds to the amino acid sequence <SEQ ID 16; ORF 004>:

1 MVERHIQHLR NGHLHLMCPS QQVRQMFGGR AYDFRADKAA GGFFGIQAHM
51 AFVHQHHAAA ALVFERYFAD DKFVGLVLRG NLRVFQTDKA DLRTGKHHAD
101 GAAPQTAADI RVAAALSPAI LPWSAASCSR GSWLDASPAM KICGILVCMV
151 SGSATGTPRA SFSILIFSKP ILSTFGRRPT AASIYSATNT PFSPSCSQWT
201 STLPSASSLT SVLASRCSFN SSPNTAFASS ETTGSEMPPM IPPKPKISTF
251 TPKRCNA*

## The following partial DNA sequence was identified in N. meningitidis <SEQ ID 17>:

ATGGTAGAAC GGCATATCCA GCATTTGCGG AACGGTCATC TTCATTTGAT 51 GTGCCCAAGC CAACAGGTGC GCCAAATGTT CGGCGGCCGG ACCTACGATT 101 TCTGCGCCGA TGAAGCGGCC GGTGGCTTTT TCGGCATACA GGCGCATATG 151 GCCTTTGTTT ACCAGCATCA CGCGGCTGCG GCCTTGGTTT TTGAACGATA CTTCGCCGAT GACAAATTCG TCGGCTTGGT ATTGCGCGGC AACCTGCGCG TATTTCAAAC CGACAAAGCC GATTTGCGGA CTGGTGAACA CTACGCCGAT 251 301 GGTGCTGCGG CGCAAACCGC CGCCGATATT CGGGTAGCGG CCGCGTTATC 351 GCCGGCAATC TTGCCTTGGT CGGCGGCTTC ATGCAGCAGG GGCAGTTGGT 401 TGGACGCGTC GCCCGCAATA AAGATATGCG GAATACTGGT CTGCATAGTC 451 AGCGGATCGG CAACGGGTAC GCCGCGCGCA TCTTTTTCGA TATTGATGTT TTCCAAACCG ATATTGTCAA CGTTCGGACG GCGGCCTACG GCTGCCAGCA 551 TATATTCGGC AACAAATACG CCTTTTTCGC CATCCTGCTC CCAATGGACT 601 TCTACATTGC CGTCTGCGTC GAGTTTGGCC TCGGTTTTAG CATCCAAATG 651 CAGTTTCAAT TCTTCACCGA ACACGGCTTT CGCCTCGTCT GAAACGACGG 701 GGTCGGAAAT GCCGCCGATG ATGCCACCCA AACCGAAAAT TTCAACTTTC 751 ACGCCCAAAC GGTGCAATGC CTGA This corresponds to the amino acid sequence <SEQ ID 18; ORF 004.a>: a004.pep

1 MVERHIQHLR NGHLHLMCPS QQVRQMFGGR TYDFCADEAA GGFFGIQAHM 51 AFVYOHHAAA ALVFERYFAD DKFVGLVLRG NLRVFQTDKA DLRTGEHYAD 101 GAAAQTAADI RVAAALSPAI LPWSAASCSR GSWLDASPAI KICGILVCIV
151 SGSATGTPRA SFSILMFSKP ILSTFGRRPT AASIYSATNT PFSPSCSQWT
201 STLPSASSLA SVLASKCSFN SSPNTAFASS ETTGSEMPPM MPPKPKISTF 251 TPKRCNA*

94.9% identity over a 257 aa overlap m004/a004

m004.pep	10 MVERHIQHLRNGHLI           MVERHIQHLRNGHLI 10	:	11111111			
m004.pep	70 ALVFERYFADDKFV0             ALVFERYFADDKFV0 70		1111111	1:1:1111	1111111111	
m004.pep	130 LPWSAASCSRGSWLI               LPWSAASCSRGSWLI 130	1111:111	1111:1111	1111111111	1.1111111	111111
m004.pep	190 AASIYSATNTPFSPS                 AASIYSATNTPFSPS 190		11111:111	1:1111111	111111111	11111
m004.pep	250 IPPKPKISTFTPKRC :           MPPKPKISTFTPKRC 250	111				

Computer analysis of this amino acid sequence gave the following results: Homology with a predicted ORF from N. gonorrhoeae

ORF 004 shows 93.4% identity over a 258 aa overlap with a predicted ORF (ORF 004.ng) from N. gonorrhoeae:

m004/g004

m004.pep	MVERH	10 IOHLRNGHIJ	20	30	40	50 GIQAHMAFVHÇ	60
g004	1111	11111111	111 1 111	3 1 1 1 1 1 1 1 1 1 1	1111111111	GIQAHMAFVHÇ             GIQAHMAFVYÇ	
		10	20	30	40	50	60
m004.pep	ALVFE	70 RYFADDKFVO	80 SLVLRGNLRV	90 FOTDKADI.RT	100	110 QTAADIRVAAA	119
g004	- 1 - 1 1	RYFADDKFVO		1     1   1   1   1   1	11111.	                 	
		70	80	90	100	110	120
m004.pep	120 ILPWS	130 AASCSRGSWI	140 DASPAMKICO	150 SILVCMVSGS	160 ATGTPRASFS	170 ILIFSKPILST	179

g004	  ILPW	IIIIIIIIII SAASCSRGSW 130	:     LDASPAIKIO  140	:	:   ATGTPRASLS 160	:         ILMFSKPILS  170	 TFGRRP 180
	180	190	200	210	220	230	239
m004.pep	TAAS	YSATNTPFS	PSCSQWTSTI	PSASSLTSVL		NTAFASSETT	CCEMDD
	111:		1111111111	111111111	1111111111		IIIIII
g004	TAAN	YSATNTPFS	PSCSQWTSTL	PSASSLTSVL	ASRCSFNSSP	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	CCEMDD
		190	200	210	220	230	240
	240	250					
m004.pep	MIPP	(PKISTFTPK	RCNAX				
	11111	11111111	1111				
g004	MIPPH	(PKISTFTPK	RCNA				
		250					

# The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 19>: g005.seq

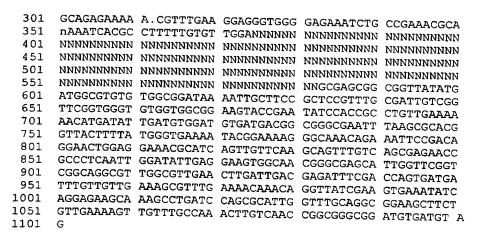
```
ATGGGGATGG ACAATATTGA TATGTTCATG CCTGAACAAG AGGAAATCCA
      ATCAATGTGG AAAGAAATTT TACTGAATTA CGGTATTTTC CTGCTCGAAC
  51
      TGCTTACCGT GTTCGGCGCA ATTGCGCTGA TTGTGTTGGC TATCGTACAG
 101
 151
      AGTAAGAAAC AGTCGGAAAG CGGCAGTGTC GTACTGACAG ATTTTTCGGA
 201 AAATTATAAA AAACAGCGGC AATCGTTTGA AACATTCTTT TTAAGCGAGG
 251 AAGAGACAAA ACATCAGGAA AAAAAAGAAA AGAAAAAGGA AAAGGCGGAA
 301 GCCAAAGCAG AGAAAAAGCG TTTGAAGGAG GGCGGGGAGA AATCTGCCGA
 351
     AACGCAAAAA TCCCGCCTTT TTGTGTTGGA TTTTGACGGC GATTTGTATG
      CACACGCCGT AGAATCCTTG CGTCATGAGA TTACGGCGGT GCTTTTGATT
 401
 451 GCCAAGCCTG AAGATGAGGT TCTGCTCAGA TTGGAAAGTC CGGGCGGCGT
 501 GGTTCACGGT TACGGTTTGG CGGCTTCGCA GCTTAGGCGT TTGCGCGAAC
 551 GCAATATTCC GCTGAccgtc gccgTCGATA AGGTCGCGGC AAGCGgcggc
 601 tatatgatgg cgtgtgtgGC GGATAAAATT GTTTCCGCtc cgtttgcggt
 651
     catcggttcg gtgggtgtgg tgGcggaagt gcCGAATATC CAccgCctGT
 701 TGAAAAAACA TGATATTGAT GTGGATGTGA TGACGGCGGG CGAATTTAAG
 751 CGCACGGTTA CTTTTATGGG TGAAAATACG GAAAAGGGCA AACAGAAATT
 801 CCGGCAGGAA CTGGAGGAAA CGCATCAGTT GTTCAAGCAG TTTGTCAGTG
 851 AAAACCGCCC CGGGTTGGAT ATTGAAAAAA TAGCGACGGG CGAGCATTGG
 901
     TTCGGCCGGC AGGCGTTGGC GTTGAACTTG ATTGACGAGA TTTCGACCAG
     TGATGATTTG TTGTTGAAAA CGTTTGAAAA CAAACAGGtt aTCGAAGTGA
 951
1001 AATATCAGGA GAAGCGAAGC CTGATCCAGC GCATTGGTTT GCAGGCGGAA
     GCTTCCGTTG AAAAGTTGTT TGCCAAACTT GTCAACCGGC GAGCGGATGT
1101 GATGTAG
```

### This corresponds to the amino acid sequence <SEQ ID 20; ORF 005.ng>:

1	MGMDNIDMFM	PEQEEIQSMW	KEILLNYGIF	LLELLTVFGA	IALIVLATVO
51	SKKQSESGSV	VLTDFSENYK	KQRQSFETFF	LSEETKHOE	KKEKKKEKAE
101	AKAEKKRLKE	GGEKSAETQK	SRLFVLDFDG	DLYAHAVESI	RHETTAVILLE
151	AKPEDEVLLR	LESPGGVVHG	YGLAASQLRR	LRERNIPLIV	AVDKVAASGG
201	YMMACVADKI	VSAPFAVIGS	VGVVAEVPNI	HRITKKHULU	UDUMTACEER
251	RTVTFMGENT	EKGKOKFROE	LEETHQLFKQ	FVSENRPGID	TEKINTCEUM
301	FGRQALALNL	IDEISTSDDL	LLKAFENKQV	TEAKAULKDS	ITOPICIONE
351	ASVEKLFAKL	VNRRADVM*		THAIRTONIA	TIĞKIĞDĞAR

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 21>: m005.seq

1	ATGGACAATA	TTGACATGTT	CATGCCTGAA	CAAGAGGAAA	ТССААТСААТ
51	GTGGAAAGAA	ATTTTACTGA	ATTACGGTAT	TTTCCTCCTC	CAACTCCTTA
101	CCGTGTTCGG	CGCAATTGCG	СТСАТТСТСТ	TEGETATECT	DCACIGCIIA
151	AAACAGTCGG	AWAGCGGCAG	TOTOTOTO	ACCCAMMENT	ACAGAGTAAG
201	TAAAAAACAG	CGGCAATCCT	TOICOIACIG	ACGGATTTT	CGGAAAATTA
251	TAAAAAACAG	CCARARCE	TTGAAGCATT	CTTTTTAAGC	GGGGAAGAGG
231	CACAACATCA	GGAAAAAGAG	GAAAAGAAAA	AGGAAAAGGC	GGAAGCCAAA



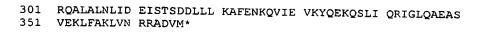
### This corresponds to the amino acid sequence <SEQ ID 22; ORF 005>:

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 23>:

```
ATGGACAATA TTGACATGTT CATGCCTGAA CAAGAGGAAA TCCAATCAAT
   1
      GTGGAAAGAA ATTTTACTGA ATTACGGTAT TTTCCTGCTC GAACTGCTTA
 101
      CCGTGTTCGG CGCAATTGCG CTGATTGTGT TGGCTATCGT ACAGAGTAAG
     AAACAGTCGG AAAGCGGCAG TGTCGTACTG ACGGATTTTT CGGAAAATTA
 151
      TAAAAAACAG CGGCAATCGT TTGAAGCATT CTTTTTAAGC GGGGAAGAGG
 201
      CAAAACATCA GGAAAAAGAG GAAAAGAAAA AGGAAAAGGC GGAAGCCAAA
 251
      GCAGAGAAAA AGCGTTTGAA GGAGGGTGGG GAGAAATCTT CCGAAACGCA
 301
      AAAATCCCGC CTTTTTGTGT TGGATTTTGA CGGCGATTTG TATGCACACG
 351
      CCGTAGAATC CTTGCGTCAT GAGATTACGG CGGTGCTTTT GATTGCCAAG
 401
      CCTGAAGATG AGGTTCTGCT TAGATTGGAA AGTCCGGGCG GCGTGGTTCA
 451
      CGGTTACGGT TTGGCGGCTT CGCAGCTTAG GCGTTTGCGC GAACGCAATA
 501
 551
      TTCCGCTGAC CGTCGCCGTC GATAAGGTGG CGGCGAGCGG TGGTTATATG
      ATGGCGTGTG TGGCGGATAA AATTGTTTCC GCTCCGTTTG CGATTGTCGG
      TTCGGTGGGT GTTGTAGCGG AAGTACCGAA TATCCACCGC CTGTTGAAAA
 651
 701 AACATGATAT TGATGTGGAT GTGATGACGG CGGGCGAATT TAAGCGCACG
     GTTACTTTTA TGGGTGAAAA TACGGAAAAG GGCAAACAGA AATTCCGACA
     GGAACTGGAG GAAACGCATC AGTTGTTCAA GCAGTTTGTC AGCGAGAACC
 801
 851
      GCCCTCAATT GGATATTGAG GAAGTGGCAA CGGGCGAGCA TTGGTTCGGT
      CGGCAGGCGT TGGCGTTGAA CTTGATTGAC GAGATTTCGA CCAGTGATGA
 901
      TTTGTTGTTG AAAGCGTTTG AAAACAAACA GGTTATCGAA GTGAAATATC
 951
     AGGAGAAGCA AAGCCTGATC CAGCGCATTG GTTTGCAGGC GGAAGCTTCT
      GTTGAAAAGT TGTTTGCCAA ACTTGTCAAC CGGCGGGCGG ATGTGATGTA
1051
1101 G
```

# This corresponds to the amino acid sequence <SEQ ID 24; ORF 005.a>:

1	MDNIDMFMPE	<b>QEEIQSMWKE</b>	ILLNYGIFLL	ELLTVFGATA	T.TVI.ATVOCK
31	VÕRERGRAAT	TDESENYKKO	ROSFEAFFLS	GEEAKHOEKE	FEFFEFFFFFF
101	AEKKKLKEGG	EKSSETQKSR	LFVLDFDGDT.	VAHAUFSTDU	ETTAILT TAIL
131	PEDEATTER	SPGGVVHGYG	LAASOLRRIR	ERNIPI TUAU	DEMANCCOAM
201	MACVADRIVS	APPAIVGSVG	VVAEVPNIHR	T.T.KKUNTDUD	TANCE CELETOR
251	VTFMGENTEK	GKQKFRQELE	ETHOLFKOFV	SENRPOLDIE	FUNTCEUMEC
				CTIVIL GUDTE	PANIGEUMER



### m005/a005 79.2% identity over a 366 aa overlap

	10	20	30	40	50	60
m005.pep	MDNIDMFMPEQEE	IQSMWKEILL	NYGIFLLELL:	<b>PVFGAIALIV</b>	LAIVQSKKQS	XSGSVVL
		111111111	1 1 1 1 1 1 3 3 3 1 4	1		111111
a005	MONTOWEMSEGEE	IOSMWKEILL	NYGIFLLELL	PVFGAIALIV	LAIVQSKKQS	ESGSVVL
	10	20	30	40	50	. 60
	70	80	90	100	110	120
m005.pep	TDFSENYKKQRQS	FEAFFLSGEE.	<b>AQHQEKEEKKI</b>	(EKAEAKAEK	KRLKEGGEKS	ΣΕΨΥΚΟΡ
		11111111	1:1111111		11111111	• 11 111
a005	TDFSENYKKQRQS	FEAFFLSGEE!	AKHQEKEEKK	EKAEAKAEKI	KRIKEGGEKS	SETOKSD
	70	80	90	100	110	120
					110	120
	130	140	150	160	170	180
m005.pep	LFVLXXXXXXXXX	XXXXXXXXXX		XXXXXXXXXX	,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,	T00
	1111	•			·AAAAAAAAA	AAAAAAA
a005	LFVLDFDGDLYAH	AVESLRHETT	AVI.I.TAKPEDE	WII DI ECDO	TOTAL DATE	
	130	140	150	160	170	
			130	100	170	180
	190	200	210	220	000	
m005.pep	XXXXXXXXXXXXX		77 DET 27 DET	ZZU	230	240
	:	IIIIIIII	VADRIASAPEA	TVGSVGVVAL	VPNIHRLLKI	KHDIDVD
a005			:	TYPOCY.C:::-	111111111	
4000	ERNIPLTVAVDKVA 190	200	210			
	130	200	210	220	230	240
	250	260	270		_0_	
m005.pep			270	280	290	300
moos.pep	VMTAGEFKRTVTFN	AGEN LEVOVOR	KEROELEETHO	LFKQFVSENF	PQLDIEEVAT	rgehwfg
a005	VMED CE EKR MYMEN	(CENTERCATOR		<u> </u>	11111111	
a003	VMTAGEFKRTVTFN 250	ZENTEKGKQF 260	(EKOETEETHÖ	LFKQFVSENF		GEHWFG
	250	260	270	280	290	300
	310	200				
m005.pep		320	330	340	350	360
moos.pep	RQALALNLIDEIST	SUULLLKAFE	NKOVIEVKYO	EKQSLIQRIG	LQAEASVEKI	FAKLVN 📑
a005	11111111111111		1111111111	11111111	111111111	11111 52
a005	ROALALNLIDEIST	SDDLLLKAFE	NKOVIEVKYO	EKQSLIQRIG	LQAEASVEKI	FAKLVN
	310	320	330	340	350	360
m005.pep	RRADVMX					
- 005						
a005	RRADVMX					

Computer analysis of this amino acid sequence gave the following results: Homology with a predicted ORF from N. gonorrhoeae

ORF 005 shows 77.0% identity over a 366 aa overlap with a predicted ORF (ORF 005.ng) from N. gonorrhoeae:

m005/g005

		10	20	30	40	50	
m005.pep	MDNI	DMFMPEQEE	IQSMWKEILL	NYGIFLLELLI	TVFGAIALIV	LAIVQSKKQS	SXSGSV
<b>q</b> 005	MCMDNT		TOSMESTATA			111111111	1
9005	MGMDNI	DIMENTEQUE	IQSMWKEILL			LAIVQSKKQS	SESGSV
		10	20	30	40	50	60
	60	70	80	90	100	110	
m005.pep	VLTDFSI	Enykkoros	FEAFFLSGEE!	AQHQEKEEKKI	(EKAEAKAEKI	KRLKEGGEKS	SAETXK
-00F	11111	1111111		::   :	11111111		1111 1
g005	VLTDFSI	ENYKKQRQS	FETFFLSEEET	rkhqekkekki	(EKAEAKAEKI	KRLKEGGEKS	SAETOK
		70	80	90	100	110	120

m005.pep		30 140 XXXXXXXXXXXX		160 XXXXXXXXX	170 XXXXXXXX	xxxxxx
g <b>00</b> 5		DLYAHAVESLRE 140	iEITAVLLIAKPI 150	EDEVLLRLESP 160	GGVVHGYGLA 170	ASQLRR 180
m005.pep		90 200 XXXXXXASGGYM	MACVADKIASAE	220 PFAIVGSVGVV	230 AEVPNIHRLL	KKHDID
g005	LRERNIPLTV 190	AVDKVAASGGYM 200	:     MACVADKIVSAE   210	PFAVIGSVGVV. 220	AEVPNIHRLL 230	 KKHDID 240
m005.pep	VDVMTAGEFK	50 260 RTVTFMGENTEK            RTVTFMGENTEK 260	GKQKFRQELEET			11111
m005.pep	FGRQALALNL	DEISTSDDLLL THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TENTS TO THE TE	f	1111-1111		
m005.pep	360 VNRRADVMX          VNRRADVMX					

# The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 25>: g006. seq

```
1 ATGCTGCTGG TGCTggaatt ttggttCGGc gtGtCGGCGG TGGGCatact
51 tgCGTTGTTT TTATGGCttt TGCCACGTTT TGCCGCCATC AGCGAAAACC
101 TGTATTTCCG CCTGAACAAC AGCTTGGAAC gcgACAACCA CTTTATCCGA
151 AAAGGCGACG AGCGGCAGCT GTACCGCCAT TACGGACTGG TTTCGCGCCT
201 GCGTGTGCTG ATTTCCAACC GCGAAGCCTT CGGCTATCTC TGCGTCGGCG
251 CGGCGATGGG TATTTTGTTC GGCTTTGCTT TTGTGATGAT GACGCTCAAA
301 GGCTACGGCA GCGCGGGCA TATTTATTCG GTCGGCACTT ATCTGTGGAT
351 GTTTGCCATG AGTTTGGACG ATGTGCCGCG ATTGGTCGAA CAATATTCCA
401 ATTTGAAAGA CATCGGACAA CGGATAGAGT GGTCGGAACG GAACATCAAA
```

# This corresponds to the amino acid sequence <SEQ ID 26; ORF 006.ng>:

- 1 MLLVLEFWFG VSAVGILALF LWLLPRFAAI SENLYFRLNN SLERDNHFIR 51 KGDERQLYRH YGLVSRLRVL ISNREAFGYL CVGAAMGILF GFAFVMMTLK 101 GYGSAGHIYS VGTYLWMFAM SLDDVPRLVE QYSNLKDIGQ RIEWSERNIK
- 151 AGT*

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 27>:

101 151 201 251	TGCGTTGTTT TGTATTTCCG AAAGGCGACC GCGTGTGCTG CGGCGATGGG	CCTGAACAAC GGCGGCAGCT ATTTCCAACC TATTTTGTTC	TGCCACGTTT AGCTTGGAAC GTACCGCCAT GCGAAGCCTT GGCTTTGCTT	TGCCGCCATC GCGACAACCA TACGGACTGC CGGCTATCTC TTGTGATGAT	AGCGAAAACC CTTTATCCGA TTGCGCGCCT TGCGTCGGCA
301	GGCTACAGCA	GCGCGGGGCA	TGTCTATTCG	GTCGGCACTT	ATCTGTGGAT
			TOTCIALICG	GICGGCACTT	ATCTGTGG



- 351 GTTTGCCATG AGTTTGGACG ACGTGCCGCG ATTGGTCGAA CAATATTCCA
- 401 ATTTGAAAGA CATCGGACAA CGGATAGAGT GGTCGGAACG GAACATCAAA
- 451 GCCGGAACTTGA

#### This corresponds to the amino acid sequence <SEQ ID 28; ORF 006>: m006.pep

- 1 MLLVLEFWVG VSAVGILALF LWLLPRFAAI SENLYFRLNN SLERDNHFIR
- 51 KGDRRQLYRH YGLLARLRVL ISNREAFGYL CVGTAMGILF GFAFVMMTLK
- 101 GYSSAGHVYS VGTYLWMFAM SLDDVPRLVE QYSNLKDIGQ RIEWSERNIK

### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 29>:

#### a006.seq

- ATGCTGCTGG TGCTGGAATT TTGGGTCGGC GTGTCGGCGG TGGGCATACT 1 51 TGCGTTGTTT TTATGGCTTT TGCCACGTTT TGCCGCCATC AGCGAAAACC
- TGTATTTCCG CCTGAAGAAC AGCTTGGAAC GCGACAACCA CTTTATCCGA 101 151 AAAGGCGACG AGCGGCAGCT GGACCGCCAT TACGGACTGC TTGCGCGCCCT
- 201 GCGTGTGCTG ATTTCCAACC GCGAAGCCTT CGGCTATCTC TGCGTCGGCA
- 251 CGGCGATGGG TATTTTGTTC GGCTTTGCTT TTGTGATGAT GACGCTCAAA
- 301 GGCTACAGCA GCGCGGGGCA TGTCTATTCG GTCGGCACTT ATCTGTGGAT
- 351 GTTTGCCATA AGTTTGGACG ACGTGCCGCG ATTGGTCGAA CAATATTCCA
- ATTTGAAAGA CATCGGACAA CGGATAGAGT GGTCGAAACG GAACATCAAA
- 451 GCCGGAACTT GA

### This corresponds to the amino acid sequence <SEQ ID 30; ORF 006.a>:

#### a006.pep

- MLLVLEFWVG VSAVGILALF LWLLPRFAAI SENLYFRLKN SLERDNHFIR
- 51 KGDERQLDRH YGLLARLRVL ISNREAFGYL CVGTAMGILF GFAFVMMTLK
- 101 GYSSAGHVYS VGTYLWMFAI SLDDVPRLVE QYSNLKDIGQ RIEWSKRNIK

#### 96.7% identity over a 153 aa overlap m006/a006

	10	20	30	40	50	60
m006.pep	MLLVLEFWVGVSAV	GILALFLWLI	LPRFAAISENI	YFRLNNSLER	DNHFIRKGDF	ROLYRH
	1111111111111		[	1111:1111	111111111:	111 11
a006	MLLVLEFWVGVSAV	GILALFLWLI	LPRFAAISENI	YFRLKNSLER	DNHFIRKGDE	ROLDRH
	10	20	30	40	50	60
	70	80	90	100	110	120
m006.pep	YGLLARLRVLISNR	EAFGYLCVG	AMGILFGFAF	VMMTLKGYSS		
	[ [ [ [ [ [ [ [ [ [ [ [ [ [ [ [ [ [ [ [	111111111		111111111		11111:
a006	YGLLARLRVLISNR	EAFGYLCVG	AMGILFGFA	VMMTLKGYSS	AGHVYSVGTY	LWMFAI
	70	80	90	100	110	120
	130	140	150			
m006.pep	SLDDVPRLVEQYSN	LKDIGQRIEV	<b>VSERNIKAGTX</b>			
	F	111111111	11:11:11:11			

150

Computer analysis of this amino acid sequence gave the following results:

SLDDVPRLVEQYSNLKDIGQRIEWSKRNIKAGTX

140

### Homology with a predicted ORF from N. gonorrhoeae

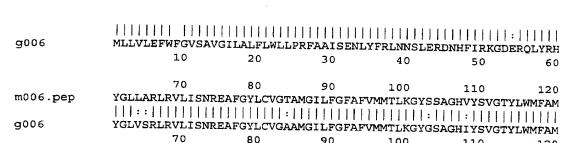
130

ORF 006 shows 95.4% identity over a 153 aa overlap with a predicted ORF (ORF 006.ng) from N. gonorrhoeae:

m006/g006

a006

10 20 30 40 60 MLLVLEFWVGVSAVGILALFLWLLPRFAAISENLYFRLNNSLERDNHFIRKGDRRQLYRH m006.pep



100

110

130 140 150 m006.pep SLDDVPRLVEQYSNLKDIGQRIEWSERNIKAGTX q118 SLDDVPRLVEQYSNLKDIGQRIEWSERNIKAGT 130 140

### The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 31>: q006-1.seq

80

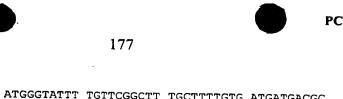
ATGTGGAAAA TGTTGAAACA CATAGCCAAA ACCCACCGCA AGCGATTGAT 1 TGGCACATTT TCCCCGGTCG GACTGGAAAA CCTTTTGATG CTGGGGTATC 51 CGGTGTTTGG CGGCTGGGCG ATTAATGCCG TGATTGCGGG GAGGGTGTGG 101 CAGGCGTTGC TGTACGCTTT GGTTGTATTT TTGATGTGGC TGGTCGGTGC GGCACGGCGG ATTGCCGATA CGCGCACGTT TACGCGGATT TATACCGAAA 201 TCGCCGTGCC GGTTGTGTTG GAACAACGGC AGCGGCAAGT CCCGCATTCA 251 GCGGTAACTG CACGGGTTGC CCTGTCGCGT GAATTTGTCA GCTTTTTTGA 301 AGAACACCTG CCGATTGCCG CGACATCCGT CGTATCCATA TTCGGCGCGT 351 GCATCATGCT GCTGGTGCTG GAATTTTGGG TCGGCGTGTC GGCGGTGGGC ATACTTGCGT TGTTTTTATG GCTTTTGCCA CGTTTTGCCG CCATCAGCGA 451 501 AAACCTGTAT TTCCGCCTGA ACAACAGCTT GGAACGCGAC AACCACTTTA TCCGAAAAGG CGACGAGCGG CAGCTGTACC GCCATTACGG ACTGGTTTCG 551 CGCCTGCGTG TGCTGATTTC CAACCGCGAA GCCTTCGGCT ATCTCTGCGT 601 CGGCGCGGCG ATGGGTATTT TGTTCGGCTT TGCTTTTGTG ATGATGACGC 651 TCAAAGGCTA CGGCAGCGCG GGGCATATTT ATTCGGTCGG CACTTATCTG 701 751 TGGATGTTTG CCATGAGTTT GGACGATGTG CCGCGATTGG TCGAACAATA TTCCAATTTG AAAGACATCG GACAACGGAT AGAGTGGTCG GAACGGAACA 851 TCAAAGCCGG AACTTGA

### This corresponds to the amino acid sequence <SEQ ID 32; ORF 006-1.ng>: g006-1.pep

MWKMLKHIAK THRKRLIGTF SPVGLENLLM LGYPVFGGWA INAVIAGRVW QALLYALVVF LMWLVGAARR IADTRTFTRI YTEIAVPVVL EQRQRQVPHS 51 AVTARVALSR EFVSFFEEHL PIAATSVVSI FGACIMLLVL EFWVGVSAVG 101 ILALFLWLLP RFAAISENLY FRLNNSLERD NHFIRKGDER QLYRHYGLVS RLRVLISNRE AFGYLCVGAA MGILFGFAFV MMTLKGYGSA GHIYSVGTYL WMFAMSLDDV PRLVEQYSNL KDIGQRIEWS ERNIKAGT*

### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 33>: m006-1.seg

ATGTGGAAAA TGTTGAAACA CATAGCCCAA ACCCACCGCA AGCGATTGAT 1 51 TGGCACATTT TCCCTGGTCG GACTGGAAAA CCTTTTGATG CTGGTGTATC CGGTGTTTGG CGGCCGGGCG ATCAATGCCG TGATTGCGGG GGAGGTGTGG CAGGCGTTGC TGTACGCTTT GGTTGTGCTT TTGATGTGGC TGGTCGGTGC GGTGCGGCGG ATTGCCGATA CGCGCACGTT TACGCGGATT TATACCGAAA TCGCCGTGCC GGTCGTGTTG GAACAGCGGC AGCGACAAGT CCCGCATTCG GCGGTAACTG CGCGGGTTGC CCTGTCGCGT GAGTTTGTCA GCTTTTTTGA 301 351 AGAACACCTG CCGATTGCCG CGACATCCGT CGTATCCATA TTCGGCGCGT GCATCATGCT GCTGGTGCTG GAATTTTGGG TCGGCGTGTC GGCGGTGGGC 401 451 ATACTTGCGT TGTTTTTATG GCTTTTGCCA CGTTTTGCCG CCATCAGCGA 501 AAACCTGTAT TTCCGCCTGA ACAACAGCTT GGAACGCGAC AACCACTTTA 551 TCCGAAAAGG CGACCGGCGG CAGCTGTACC GCCATTACGG ACTGCTTGCG 601 CGCCTGCGTG TGCTGATTTC CAACCGCGAA GCCTTCGGCT ATCTCTGCGT



```
651 CGGCACGGCG ATGGGTATTT TGTTCGGCTT TGCTTTTGTG ATGATGACGC
701 TCAAAGGCTA CAGCAGCGCG GGGCATGTCT ATTCGGTCGG CACTTATCTG
    TGGATGTTTG CCATGAGTTT GGACGACGTG CCGCGATTGG TCGAACAATA
    TTCCAATTTG AAAGACATCG GACAACGGAT AGAGTGGTCG GAACGGAACA
851 TCAAAGCCGG AACTTGA
```

### This corresponds to the amino acid sequence <SEQ ID 34; ORF 006-1>:

m006-1.pep

MWKMLKHIAQ THRKRLIGTF SLVGLENLLM LVYPVFGGRA INAVIAGEVW QALLYALVVL LMWLVGAVRR IADTRTFTRI YTEIAVPVVL EQRQRQVPHS 51 101 AVTARVALSR EFVSFFEEHL PIAATSVVSI FGACIMLLVL EFWVGVSAVG ILALFLWLLP RFAAISENLY FRLNNSLERD NHFIRKGDRR QLYRHYGLLA 201 RLRVLISNRE AFGYLCVGTA MGILFGFAFV MMTLKGYSSA GHVYSVGTYL

251 WMFAMSLDDV PRLVEQYSNL KDIGQRIEWS ERNIKAGT*

#### 95.5% identity in 288 aa overlap m006-1/g006-1

m006-1.pep	10 MWKMLKHIAQTHRF	20 CRLIGTFSLVO	30 SLENLLMLVY		50 /IAGEVWQAL	
g006-1	MWKMLKHIAKTHRE	RLIGTFSPV	LENLIMICY	PVFGGWATNAN	11111111111111111111111111111111111111	111111. TVATS75
-	10	20	30	40	50	60
				40	30	60
	70	80	90	100	110	120
m006-1.pep	LMWLVGAVRRIADI	RTFTRTYTET			יים מותים דענות! בדים	120
• •		111111111		TITLITITI SVÖAEUSMAT	KVALSKEFV	SEFEERL
g006-1	LMWLVGAARRIADT	፣ ብጥሃ T ዓጥዣ ም <b>ያ</b>	ANDINITEDO	11111111111111111111111111111111111111		111111
<b>9</b>	70	80	90	100		
	. 0	00	90	100	110	120
	130	140	150	160	170	100
m006-1.pep	PIAATSVVSIFGAC				170	180
more arpor		11111111	GVSAVGILAI	CETMTTEKEVE	TSENLYFRL	NNSLERD
g006-1	PIAATSVVSTECAC	IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII		[	7	111111
9000 1	PIAATSVVSIFGAC	140				
	130	140	150	160	170	180
	100					
	190	200	210	220	230	240
m006-1.pep	NHFIRKGDRRQLYR	HYGLLARLRV	LISNREAFGY	LCVGTAMGIL	FGFAFVMMT:	LKGYSSA
			1111111111			:
g006-1	NHFIRKGDERQLYR	HYGLVSRLRV	LISNREAFGY	/LCVGAAMGII	FGFAFVMMT	LKGYGSA
	190	200	210	220	230	240
	250	260	270	280	289	
m006-1.pep	GHVYSVGTYLWMFA	MSLDDVPRLV	EQYSNLKDIG	ORIEWSERNI	KAGTX	
	11:1111111111	1111111111	111111111		11111	
g006-1	GHIYSVGTYLWMFA	MSLDDVPRLV	EOYSNLKDTO	ORIEWSERNT	KACTY	
	250	260	270	280	101017	
				200		

### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 35>: a006-1.seg (partial)

oo-r.seq	(partial)				
1.	.AGCCAAAACC	ACCGCAAGCG	ATTGATTGGC	ACATTTTTC	TGGTCGGACT
51	GGAAAACCTT	TTGATGCTGG	TGTATCCGGT	GTTTGGCGGC	TGGGCGATTA
101	ATGCCGTGAT	TGCGGGGCAG	GCGTGGCAGG	CGTTGCTGTA	CGCTTTGGTT
151	GTGCTTTTGA	TGTGGCTGGT	CGGTGCGGCG	CGGCGGATTG	CCGATACGCG
201	CACGTTTACG	CGGATTTATA	CCGAAATCGC	CGTGCCGGTT	GTGTTGGAAC
251	AGCGGCAGCG	GCAAGTCCCG	CATTCGGCGG	TAACTGCGCG	GGTTGCCCTG
301	TCGCGTGAGT	TTGTCAGCTT	TTTTGAAGAA	CACCTGCCGA	TTGCCGCGAC
351	ATCCGTCGTA	TCCATATTCG	GCGCGTGCAT	CATGCTGCTG	GTGCTGGAAT
401	TTTGGGTCGG	CGTGTCGGCG	GTGGGCATAC	TTGCGTTGTT	TTTATCCCTT
451	TTGCCACGTT	TTGCCGCCAT	CAGCGAAAAC	CTGTATTTCC	GCCTGAAGAA
501	CAGCTTGGAA	CGCGACAACC	ACTTTATCCG	AAAAGGCGAC	GAGCGGCAGC
551	TGGACCGCCA	TTACGGACTG	CTTGCGCGCC	TGCGTGTGCT	GATTTCCAAC
601	CGCGAAGCCT	TCGGCTATCT	CTGCGTCGGC	ACGGCGATGG	CTATTCCAAC
651	CGGCTTTGCT	TTTGTGATGA	TGACGCTCAA	AGGCTACAGC	ACCCCCCCC
					<b>プログログログログ</b>





701	ATGTCTATTC GG	TCGGCACT T	ATCTGTGGA	TGTTTGCCAT	AAGTTTGGAG	-
<b>7</b> 51	GACGTGCCGC GA	TTGGTCGA A	CAATATTCC	AATTTCAAAC	ACATCCCACA	`
801	ACGGATAGAG TO	GTCGAAAC G	CAACATCAA	ACCCCCAACT	MCA CGGACE	4
			or a rotation of	AGCCGGAACI	IGA	
This corresponds	to the amino aci	d seguence	CEO ID	26. ODE 000	. 1	
Tims corresponds	to the animo aci	u scquence	SEQ ID.	o; OKF 000	)-1.a>:	
a006-1.pep	(partial)					
1 .	.SQNHRKRLIG TF	FLVGLENL L	MLVYPVFGG	WAINAVIAGO	AWOALLYAL	7
51	VLLMWLVGAA RR	IADTRTFT R	IYTEIAVPV	VLEOROROVP	HSAVTARVAT	-
101	SREFVSFFEE HL	PIAATSVV S	IFGACIMLL	VLEFWVGVSA	VCTTALET WIT	
151	LPRFAAISEN LY	FRLKNSLE R	DNHFIRKGD	EROLDRHYGI.	TARTRUTTON	7
201	REAFGYLCVG TA	MGILFGFA F	VMMTLKGYS	SACHUVSUCT	ALMMEN TOTAL	
251	DVPRLVEQYS NL	KDIGORIE W	SKRNIKACT	*	ITMMENISTI	,
		DIOQNID W	DIMINITINGS	-		
a006-1/m006-1	95.7% identity	v in 280 a	a overlan			
	TOTAL EGGINGE	10	20	30		
a006-1.pep	COMP			30	40	50
doco i.pep	JONN	VVVTTGIEET	A CTENTTWIA	YPVFGGWAIN	VIAGQAWQAL	LYALVVL
m006-1	MINISTER TEXT TO OTHER	11111111		11111 1111	1111::111	111111
m006-1	MWKMLKHIAQTH	RKRLIGTFSL	VGLENLLMLV	YPVFGGRAIN <i>A</i>	VIAGEVWQAL	LYALVVL
	10	20	30	40	50	60
	60	70	80	90	100	110
a006-1.pep	LMWLVGAARRIA	DTRTFTRIYT	EIAVPVVLEO	ROROVPHSAVT	ARVAL SDEEN	CEEEE
		111111	1	1111111111	1111111111	111111
m006-1	LMWLVGAVRRIA	OTRTFTRIYT	EIAVPVVIEO	ROROVDHSAVT	ווווווווווו	CEEEE
	70	80	90	100		
	. •	00	20	100	110	120
	120	130	140	150		
a006-1.pep				150	160	170
dood 1.pep	PIAATSVVSIFG	ZCTMPTAPE!	WVGVSAVGIL	ALFLWLLPRFA	AISENLYFRL	KNSLERD
m006-1	DIADEGRACIO			111111111	11111111	: ! ! ! ! ! !
M008-1	PIAATSVVSIFG	ACIMULTALEFY	WVGVSAVGIL	ALFLWLLPRFA	AISENLYFRL:	NNSLERD
	130	140	150	160	170	180
	180	190	200	210	220	230
a006-1.pep	NHFIRKGDERQLI	RHYGLLARL	RVLISNREAF	GYLCVGTAMGI	LFGFA FVMMT	LKCVCCA
	11111111111			111111111	LICE FILLIER	111111
m006-1	NHFIRKGDRRQLY	RHYGLLARLE	RVLISNREAF	GYT.CVGTAMCT	T.ECETAETANAMIT	I I I I I I I I
	190	200	210	220	230	
			210	220	230	240
	240	250	260	270	200	
a006-1.pep	GHVYSVGTYLWM		ZUU TURAVENIT PA	270	280	
2000 1.рср	IIIIIIIIIIIII	TATOLDOV PRI	A PÕI ON PKD.	IGORIEWSKRN	IKAGTX	
m006-1	CHYVCYCTYLWM				11111	
111000-1	GHVYSVGTYLWME	WW2TDDA BKI			IKAGTX	
	250	260	270	280		

### The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 37>: g007.seq

1	atgaACACAA	CCCGACTGCC	GACCGCCTTC	ATCTTGTGCT	GCCTCTGcgC
51	CGCcGCTTCT	GCCGccgaca	acAGCatcat	gaCaAAAGGG	CAAAAAGTGT
101	ACGAATCcAa	CtGCATCGCC	TGCCACGGCA	AGAAAGGGGA	AGGGCGCGC
151	ACTGCGtTTC	CTccgctTTT	CCggtcgGac	totattatoa	20000000000
201	cgTCCtgctg	cacagcatgg	tcaaaggcAt	cascadaca	ttcaaagtgg
251	agcggcaaaa	cctacgacgg	atttatgCcc	acaaccacca	tcaaagtgg
301	GGACATTGCC	GCCGTCGCCA	CTTATATCAT	GAACGCCTTT	GA

### This corresponds to the amino acid sequence <SEQ ID 38; ORF 007.ng>: g007.pep

- 1 MNTTRLPTAF ILCCLCAAAS AADNSIMTKG QKVYESNCIA CHGKKGEGRG 51 TAFPPLFRSD CIMNKPHVLL HSMVKGIDGT FKVERQNLRR IYARNRHQRC
- 101 GHCRRRHLYH ERL*

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 39>:



	·					
m007.seq						
1	ATGAACACAA CCCGACTGCC GACCGCCCTC GTCTTGGGCT GCTTCTGCGC					
51	CGCCGCTTCT GCCGCCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT					
101	ACGAATCCAA CTGCGTCGCC TGCCACGGCA AAAAGGGCGA AGGCCGCGGA					
151	ACCATGTTTC CGCCGCTCTA CCGCTCCGAC TTCATCATGA AAAAACCGCA					
201	GGTGCTGCTG CACAGCATGG TCAAAGGCAT CAACGGTACA ATCAAAGTC.					
251	AACGGCAAAA CCTACAACGG ATTCATGCCC GCAACCGCCA TCAGCGATGC					
301	GGACATTGCC GCCGTCGCCA CTTATATCAT GAACGCCTTT GA					
	GAAGGCCIII GA					
COL '						
	s to the amino acid sequence <seq 007="" 40;="" id="" orf="">:</seq>					
m007.pep						
1	MNTTRLPTAL VLGCFCAAAS AADNSIMTKG QKVYESNCVA CHGKKGEGRG					
51	TMFPPLYRSD FIMKKPQVLL HSMVKGINGT IKVXRQNLQR IHARNRHQRC					
101	GHCRRRHLYH ERL*					
	·					
Th - 6-11	antial DNIA community of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the st					
The following p	artial DNA sequence was identified in N. meningitidis <seq 41="" id="">:</seq>					
a007.seq	-					
1	ATGAACACAA CCCGACTGCC GACCGCCCTC GTCTTGGGCT GCCTCTGCGC					
51	CGCCGCTTCT GCCGCCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT					
101	ACGAATCCAA CTGCGTCGCC TGCCACGGCA AAAAGGGCCGA AGGCCGCGGA					
151	ACCATGTTTC CGCCGCTCTA CCGCTCCGAC TTCATCATGA AAAAACCGCA					
201	GGTGCTGCTG CACAGCATGG TCAAAGGCAT CAACGGTACA ATCAAAGTC.					
251	AACGGCAAAA CCTACAACGG ATTCATGCCC GCCACTGCCA TCAGCGATGC					
301	GGACATTGCC GCCGTCGCCA CTTATATCAT GAACGCCTTT GA					
301	GGACATIGCC GCCGICGCCA CITATATCAT GAACGCCTTT GA					
This samesand	a to the amine anid account of ODO ID to ODD and					
	s to the amino acid sequence <seq 007.a="" 42;="" id="" orf="">:</seq>					
a007.pep						
1	MNTTRLPTAL VLGCLCAAAS AADNSIMTKG QKVYESNCVA CHGKKGEGRG					
51	TMFPPLYRSD FIMKKPQVLL HSMVKGINGT IKVXRQNLQR IHARHCHQRC					
101	GHCRRRHLYH ERL*					
m007/a007 9	7.3% identity over a 113 aa overlap					
	to restantly ever a ris an everlap					
	10 20 30 40 50 60					
m007.pep						
moo7.pep	MNTTRLPTALVLGCFCAAASAADNSIMTKGQKVYESNCVACHGKKGEGRGTMFPPLYRSD					
- 007						
a007	MNTTRLPTALVLGCLCAAASAADNSIMTKGQKVYESNCVACHGKKGEGRGTMFPPLYRSD					
	10 20 30 40 50 60					
	70 80 90 100 110					
m007.pep	FIMKKPQVLLHSMVKGINGTIKVXRQNLQRIHARNRHQRCGHCRRRHLYHERLX					
a007	FIMKKPQVLLHSMVKGINGTIKVXRQNLQRIHARHCHQRCGHCRRRHLYHERLX					
	70 80 90 100 110					
Computer analy	is of this amino poid some provides C. II					
Computer analys	is of this amino acid sequence gave the following results:					
Homology with	a predicted ORF from N. gonorrhoeae					
ORF 007 shows	86.7% identity over a 113 aa overlap with a predicted ORF (ORF 007.ng					
of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of th						
Irom IV. gonorrh	from N. gonorrhoeae:					
m007/g007	•					
-	$\alpha$					

FIMKKPQVLLHSMVKGINGTIKVXRQNLQRIHARNRHQRCGHCRRRHLYHERLX

MNTTRLPTALVLGCFCAAASAADNSIMTKGQKVYESNCVACHGKKGEGRGTMFPPLYRSD 

BNSDOCID:	<wo_< td=""><td>9957280A2</td><td>: 1 &gt;</td></wo_<>	9957280A2	: 1 >

m007.pep

m007.pep

g007



```
g007
                    CIMNKPHVLLHSMVKGIDGTFKVERQNLRRIYARNRHQRCGHCRRRHLYHERL
                                      80
                                                90
                                                        100
 The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 43>:
       g007-1.seq
                   (partial)
                ATGAACACAA CCCGACTGCC GACCGCCTTC ATCTTGTGCT GCCTCTGCGC
                CGCCGCTTCT GCCGCCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT
             51
           101 ACGAATCCAA CTGCATCGCC TGCCACGGCA AGAAAGGGGA AGGGCGCGGC
           151 ACTGCGTTTC CTCCGCTTTT CCGGTCGGAC TATATTATGA ACAAACCGCA
           201 CGTCCTGCTG CACAGCATGG TCAAAGGCAT CAACGGTACA ATCAAAGTCA
                ACGGCAAAAC CTACAACGGA TTCATGCCCG CAACCGCCAT CAGCGATGCG
                GACATTGCCG CCGTCGCCAC TTATATCATG AACGCCTTTG ACAACGGCGG
           301
           351 CGGAAGCGTT ACCGAAAAAG ACGTAAAACA GGCAAAAGGC AAAAAAAAC.
 This corresponds to the amino acid sequence <SEQ ID 44; ORF 007-1.ng>:
      g007-1.pep (partial)
            1 MNTTRLPTAF ILCCLCAAAS AADNSIMTKG QKVYESNCIA CHGKKGEGRG
51 TAFPPLFRSD YIMNKPHVLL HSMVKGINGT IKVNGKTYNG FMPATAISDA
           101 DIAAVATYIM NAFDNGGGSV TEKDVKQAKG KKN...
 The following partial DNA sequence was identified in N. meningitidis <SEQ ID 45>:
      m007-1.seq
               ATGAACACAA CCCGACTGCC GACCGCCCTC GTCTTGGGCT GCTTCTGCGC
             1
                CGCCGCTTCT GCCGCCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT
            51
           101 ACGAATCCAA CTGCGTCGCC TGCCACGGCA AAAAGGGCGA AGGCCGCGGA
           151 ACCATGTTTC CGCCGCTCTA CCGCTCCGAC TTCATCATGA AAAAACCGCA
                GGTGCTGCTG CACAGCATGG TCAAAGGCAT CAACGGTACA ATCAAAGTCA
                ACGGCAAAAC CTACAACGGA TTCATGCCCG CAACCGCCAT CAGCGATGCG
           251
                GACATTGCCG CCGTCGCCAC TTATATCATG AACGCCTTTG ACAACGGCGG
           301
                CGGAAGCGTT ACCGAAAAAG ACGTAAAACA GGCAAAAAGC AAAAAAAACT
           351
           401 AA
This corresponds to the amino acid sequence <SEQ ID 46; ORF 007-1>
     m007-1.pep
               MNTTRLPTAL VLGCFCAAAS AADNSIMTKG QKVYESNCVA CHGKKGEGRG
               TMFPPLYRSD FIMKKPOVLL HSMVKGINGT IKVNGKTYNG FMPATAISDA
           101 DIAAVATYIM NAFDNGGGSV TEKDVKQAKS KKN*
m007-1 / g007-1 91.7% identity in 133 aa overlap
                                    20
                                              30
                                                        40
                  {\tt MNTTRLPTALVLGCFCAAASAADNSIMTKGQKVYESNCVACHGKKGEGRGTMFPPLYRSD}
     m007-1.pep
                  MNTTRLPTAFILCCLCAAASAADNSIMTKGQKVYESNCIACHGKKGEGRGTAFPPLFRSD
     g007-1
                          10
                                              30
                                                        40
                                                                  50
                          70
                                    80
                                              90
                                                       100
                                                                 110
                  FIMKKPQVLLHSMVKGINGTIKVNGKTYNGFMPATAISDADIAAVATYIMNAFDNGGGSV
     m007-1.pep
                  YIMNKPHVLLHSMVKGINGTIKVNGKTYNGFMPATAISDADIAAVATYIMNAFDNGGGSV
     g007-1
                          70
                                    80
                                              90
                                                       100
                         130
                  TEKDVKQAKSKKNX
     m007-1.pep
                  111111111:111
     q007-1
                  TEKDVKQAKGKKN
                         130
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 47>:
     a007-1.seg
                 (partial)
              ATGAACACAA CCCGACTGCC GACCGCCCTC GTCTTGGGCT GCCTCTGCGC
```

51 CGCCGCTTCT GCCGCCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT 101 ACGAATCCAA CTGCGTCGCC TGCCACGGCA AAAAGGGCCGA AGGCCGCGGA

151	ACCATGTTTC	CGCCGCTCTA	CCGCTCCGAC	TTCATCATGA	AAAAACCGCA
201	GGTGCTGCTG	CACAGCATGG	TCAAAGGCAT	CAACGGTACA	ATCAAAGTCA
251	ACGGCAAAAC	CTACAACGGA	TTCATGCCCG	CCACTGCCAT	CAGCGATGCG
301	GACATTGCCG	CCGTCGCCAC	TTATATCATG	AACGCCTTTG	ACAACGGCGG
351	CGGAAGCGTT	ACCGAAAAAG	ACGTAAAACA	GGCAAAAAAC	AAAAAA

## This corresponds to the amino acid sequence <SEQ ID 48; ORF 007-1.a>: a007-1.pep (partial)

- 1 MNTTRLPTAL VLGCLCAAAS AADNSIMTKG QKVYESNCVA CHGKKGEGRG
- 51 TMFPPLYRSD FIMKKPQVLL HSMVKGINGT IKVNGKTYNG FMPATAISDA
- 101 DIAAVATYIM NAFDNGGGSV TEKDVKQAKN KK..

### m007-1/a007-1 98.5% identity in 132 aa overlap

	10	20	30	40	50	60
m007-1.pep	MNTTRLPTALVLGCF	CAAASAAD	NSIMTKGQKVY	ESNCVACHGE	KGEGRGTMF	PPLYRSD
		11111111		11111111111	11111111	
a007-1	MNTTRLPTALVLGCL	CAAASAAD	NSIMTKGQKVY	ESNCVACHGE	KGEGRGTMF	PPLYRSD
	10	20	30	40	50	60
	70	80	90	100	110	120
m007-1.pep	FIMKKPQVLLHSMVK	GINGTIKV	NGKTYNGFMPA [*]	TAISDADIAA	VATYIMNAFI	ONGGGSV
	111111111111111	HHHHH	HIHIHIH	111111111	111111111	
a007-1	FIMKKPQVLLHSMVK	GINGTIKV	NGKTYNGFMPA'	TAISDADIA	VATYIMNAFI	NGGGSV
	70	80	90	100	110	120
	130					
m007-1.pep	TEKDVKQAKSKKNX					
a007-1	TEKDVKQAKNKK					
	130					

# The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 49>: g008.seq

```
ATGAACAACA GACATTTGC CGTCAtcgCC TTGGGCAGCA ACCTTGACAA
CCCCGCACAA CAAATacgcg gcgcattaga cgcgctctcg tcccatcctg
acatccggct tgaaCaggtt tcctcactgt aTatgaccgc acctgtcggt
tacgAcaaTC agcccgATTT CATCaatgcc gTCTgcaccg TTTCCACCAC
CtTGGACGGC ATTGcccTGC TTGCCgaACT CAAccgTATC GAAGCCGATT
CGCACGCGA CGCAGATTC CGCAATGCAC CGCGCACATT GGATTTGGAC
ATTATCGACT TTGACGGCAT CTCCAGCGAC GACCCCCGCC TTACCCTGCC
GCATCCGCGC GCGCACGAAC GCAGTTTCGT CATACGCCCT TTGGCAGAAA
CTCCCCTGA TTTTATTTTG GGAAAATACG GAAAGGTTGT CGAATTGTCA
AAACGGCTGG GCAATCAAGG CATCCGTCTT TTACCGGACA GGTAA
```

## This corresponds to the amino acid sequence <SEQ ID 50; ORF 008.ng>: g008.pep

1 MNNRHFAVIA LGSNLDNPAQ QIRGALDALS SHPDIRLEQV SSLYMTAPVG 51 YDNQPDFINA VCTVSTTLDG IALLAELNRI EADFGRERSF RNAPRTLDLD 101 IIDFDGISSD DPRLTLPHPR AHERSFVIRP LAEILPDFIL GKYGKVVELS 151 KRLGNQGIRL LPDR*

## The following partial DNA sequence was identified in N. meningitidis <SEQ ID 51>: m008.seq

1	ATGAACAACA	GACATTTTGC	CGTCATCGCC	CTGGGCAGTA	ATCTTGAAAA
51	CCCTGCTCAA	CAGGTACGCG	CCGCATTGGA	CACGCTGTCG	TCCCATCCTG
101	ACATCCGTCT	TAAACAGGCT	TCCTCACTGT	ATATGACCGC	GCCCGTCGGT
151	TACGACAATC	AGCCCGATTT	TGTCAATGCC	GTCTGCACCG	TTTCCACCAC
201	TCTGGACGGC	ATTGCCYTGC	TTGCCGAACT	CAACCGTATC	GAGGCTGATT
251	TCGGACGCGA	ACGCAGCTTC	CGCAACGCGC	CGCGCACATT	GKATTTGGAC
301	ATTATCGACT	TTGACGGCAT	CTCCAGCGAC	GACACSCGAC	TcACCtTGCC

351 401 451	GCATCCGCGC GCGCACGAAC GCAGTTTCGT CATCCGCCCT TTGGCAGAAA TCCTCCCTGA TTTTGTTTTA GGAAAACACG GAAAGGTTGC CGAATTGTCA AAACGGYTGG GCAATCAAGG TATCCGTCTT TTACCGGACA GGTAATT			
This corresponds	s to the amino acid sequence <seq 008="" 52;="" id="" orf="">:</seq>			
1 51 101	MNNRHFAVIA LGSNLENPAQ QVRAALDTLS SHPDIRLKQA SSLYMTAPVG YDNQPDFVNA VCTVSTTLDG IALLAELNRI EADFGRERSF RNAPRTLXLD IIDFDGISSD DTRLTLPHPR AHERSFVIRP LAEILPDFVL GKHGKVAELS KRLGNQGIRL LPDR*			
The following pa	rtial DNA sequence was identified in N. meningitidis <seq 53="" id="">:</seq>			
1 51 101 151 201 251 301 351 0	ATGAACAACA GACATTTTGC CGTCATCGCC CTGGGCAGTA ATCTTGAAAA CCCTGCCCAA CAGGTACGCG CCGCATTGGA CACGCTGTCG TCCCATCCTG ACATCCGTCT TAAACAGGCT TCCTCACTGT ATATGACCGC GCCCGTCGGT TACGACAATC CGTCAATGCC GTCTGCACCG TTTCCACCAC CTTGGACGGC ATTGCCCTGC TGCCGAACT CAACCGTATC GAAGCCGATT TCGGACGCGA ACGCAGCTC CGCAACGCGC CGCGCACATT GGATTTGGAC ATTATCGACT TTGACGGCAT CTCCAGCGAC GACCCCCGAC TCACCCTGCC GCATCCCCCGA TTTTATTTTTG GGAAAACACCG GAAAGGTTGC AAACGGCTGG GCAATGAACACG GAAAGGTTGC CGAATTGTCA			
This corresponds	to the amino acid sequence <seq 008.a="" 54;="" id="" orf="">:</seq>			
a008.pep 1 N 51 N 101 D 151 F	MNNRHFAVIA LGSNLENPAQ QVRAALDTLS SHPDIRLKQA SSLYMTAPVG VDNQPDFVNA VCTVSTTLDG IALLAELNRI EADFGRERSF RNAPRTLDLD UIDFDGISSD DPRLTLPHPR AHERSFVIRP LAEILPDFIL GKHGKVAELS KRLGNQGIRL LPDK*  6% identity over a 164 aa overlap			
	10 20 30 40 50 60			
m008.pep	MNNRHFAVIALGSNLENPAQQVRAALDTLSSHPDIRLKQASSLYMTAPVGYDNQPDFVNA			
m008.pep	70 80 90 100 110 120 VCTVSTTLDGIALLAELNRIEADFGRERSFRNAPRTLXLDIIDFDGISSDDTRLTLPHPR			
a008	VCTVSTTLDGIALLAELNRIEADFGRERSFRNAPRTLDLDIIDFDGISSDDPRLTLPHPR 70 80 90 100 110 120			
m008.pep	130 140 150 160 AHERSFVIRPLAEILPDFVLGKHGKVAELSKRLGNQGIRLLPDRX			
Computer analysis of this amino acid sequence gave the following results:  Homology with a predicted ORF from N. gonorrhoeae  ORF 008 shows 92.7% identity over a 164 aa overlap with a predicted ORF (ORF008.ng) from N. gonorrhoeae:  m008/g008				
m008.pep	10 20 30 40 50 60.  MNNRHFAVIALGSNLENPAQQVRAALDTLSSHPDIRLKQASSLYMTAPVGYDNQPDFVNA			

عالم الموادات الأميلي الماليان وتران

	10	20	30	40	50	60
m008.pep	70 VCTVSTTLDGIALLA	80 AFINRTEADE	90	100	110	120
шоострор					DG155DDTRI	TLPHPR
g008	VCTVSTTLDGIALLA	AELNRIEADE			DGISSDDPRI	TLPHPR
	70	80	90	100	110	120
	130	140	150	160		
m008.pep	AHERSFVIRPLAEII	LPDFVLGKHG	KVAELSKRLG	NQGIRLLPDR	x	
			11:111111	1111111111	1	
g008	g008 AHERSFVIRPLAEILPDFILGKYGKVVELSKRLGNQGIRLLPDRX					
	130	140	150	160		

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 55>: g009.seq

- 51 CGAACAAAAT ACCCATCGCC GCGCCGACGC AGAGATAGCC GAAGGCTTCG 101 CGGTTGGAAA TCAGCACACG CAGGCGCGAA ACCAGTCCGT AATGGCGGTA
- 151 CAGCTGCCGC TCGTCGCCTT TTCGGATAAA GTGGTTGTCG CGTTCCAAGC
  201 TGTTGTTCAG GCGGAAATAC AGGTTTTCGC TGATGGCGGC AAAACGTGGC
  251 AaaaGCCATA A

This corresponds to the amino acid sequence <SEQ ID 56; ORF 009.ng>: g009.pep

- MPRAAVAFER HHHKSKAEQN THRRADAEIA EGFAVGNQHT QARNQSVMAV
- 51 QLPLVAFSDK <u>VVVAFQAVVQ</u> AEIQVFADGG KTWQKP*

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 57>: m009.seq

- 51 CGAACAAAAT ACCCATCGCC GTGCCGACGC AGAGATAGCC GAAGGCTTCG 101 CGGTTGGAAA TCAGCACACG CAGGCGCGCA AGCAGTCCGT AATGGCGGTA
- 151 CAGCTGCCGC CGGTCGCCTT TTCGGATAAA GTGGTTGTCG CGTTCCAAGC
- 201 TGTTGTTCAG GCGGAAATAC AGGTTTTCGC TGATGGCGGC AAAACGTGGC
- 251 AAAAGCCATA A

This corresponds to the amino acid sequence <SEQ ID 58; ORF 009>: m009.pep

- MPRAAVAFER HHHKSKAEQN THRRADAEIA EGFAVGNQHT QARKQSVMAV
- 51 QLPPVAFSDK VVVAFQAVVQ AEIQVFADGG KTWQKP*

Computer analysis of this amino acid sequence gave the following results: Homology with a predicted ORF from N. gonorrhoeae

ORF 009 shows 97.7% identity over a 86 aa overlap with a predicted ORF (ORF 009.ng) from N. gonorrhoeae:

m009/g009

	10	20	30	40	50	60
m009.pep	MPRAAVAFERHHHKS	KAEQNTHRE	RADAEIAEGFA	VGNQHTQARK	QSVMAVOLPP'	VAFSDK
				111111111:	111111111	11111
g009	MPRAAVAFERHHHKS	SKAEQNTHRE	RADAEIAEGFA	VGNQHTQARN	QSVMAVQLPL	VAFSDK
	10	20	30	40	50	60
	70	80				
m009.pep	VVVAFQAVVQAEIQV	FADGGKTWC	KPX			
	1111111111111111	111111111	111			
g009	VVVAFQAVVQAEIQV	FADGGKTWC	KPX			
	70	80				

```
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 59>:
     a009.seq
             CGAACAAAAT ACCCATCGCC GTGCCGACGC AGAGATAGCC GAAGGCTTCG
          51
         101 CGGTTGGAAA TCAGCACACG CAGGCGCGCA AGCAGTCCGT AATGGCGGTC
         151 CAGCTGCCGC TCGTCGCCTT TTCGGATAAA GTGGTTGTCG CGTTCCAAGC
         201 TGTTCTTCAG GCGGAAATAC AGGTTTTCGC TGATGGCGGC AAAACGTGGC
         251 AAAAGCCATA A
This corresponds to the amino acid sequence <SEQ ID 60; ORF 009.a>:
    a009.pep
```

MPRAAVAFER HHHKSKAEQN THRRADAEIA EGFAVGNQHT QARKQSVMAV 1

QLPLVAFSDK VVVAFQAVLQ AEIQVFADGG KTWQKP* 51

### m009/a009 97.7% identity over a 86 aa overlap

```
40
          MPRAAVAFERHHHKSKAEQNTHRRADAE IAEGFAVGNQHTQARKQSVMAVQLPPVAFSDK
m009.pep
          MPRAAVAFERHHKSKAEQNTHRRADAEIAEGFAVGNQHTQARKQSVMAVQLPLVAFSDK
a009
                                 30
                                         40
                                                  50
                 70
                         80
m009.pep
          VVVAFQAVVQAEIQVFADGGKTWQKPX
          1111111:111111111111111111
a009
          VVVAFQAVLQAEIQVFADGGKTWQKPX
                 70
                         80
```

### The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 61>: g010.seq

1	ATGGGTTTTC	CTGTTCGCAA	GTTTGATGCC	GTGATTGTCG	GCGGTGGCGG
51	TGCAGGTTTA	CGTGCAGCCC	TCCAATTATC	CAAATCCGGT	
101	CCGTTTTGTC	TAAAGTGTTC	CCGACCCGCT	CGCATACCGT	
151	GGCGGTATTT	CCGCCTCTCT	GGGTAATGTG	CAGGAGGACC	-100000CAG
201	GCACATGTAC		AAGGTTCCGA		
251	CGATTGAGTT		GCTGCGCCTG		or 10 OI II 10 II 1 G
301	CACATGGGTA	TGCCTTTTGA	CCGCGTTGAA	AGCGGCAAAA	TGAGTTGGAA
351	TCCTTTCGGC		CCGAACATGG		
401	CATGTGCGGT	TGCCGACCGT		CGATGTTGCA	GTAGAACGTG
451	CAACAAAACG			TTTGTGGAAT	
501	AGATTTGATT		ACGGCGATGT	CGTCGGCGTA	GGACGGCGCA
551	AAATGGAAAC	GGGCGAAGTT	TATATTTTCC	ACGCCAAGGC	
601	GCTACCGGTG	GCGGCGGTCG	TATTTATGCT		CGTGATGTTT
651	GAATACCGGT	GACGGTTTGG	GCATTTGCGC	CCGTGCGGGC	ATGCTTATAT
701	AAGATATGGA				ATTCCGTTGG
751	GTGTTGATTA	CCGAAGGCGT			GGGTGCGGGC
801	cgacggcgaA	CGCTTTATGG	AAcastatas		TGTTGAacgc
851	CTTCTCGCga	cgtGGTTTCA	Cacacatac		aAagaCTTGG
901	cgcggctqTG	GtaaAAAcaA	agaCCacatC	CGatggaAAt	ctatgaaggt
951			AAAAACTGCC		TCGACcAtAt
1001	TTCagtttgc		CCGATTAAAG		GAGATTTCCA
1051	ACTACCCACT	ATATGATGGG	CGGCATTCcg	ACCCGATTCC	ggttgTGCCG
1101	TGTTGTTCCG		AGTACGAAGT		ACGGTGAAGT
1151	CCGCAGGTGA	GTGCGCCTCT	GCTTCCGTAC		GGCCTGTATG
1201	ACGAACTCCC				CCGTTTGGGT
		- 00100AC11	GGTGGTGTTC	cgcccaaccc	cccggtga

This corresponds to the amino acid sequence <SEQ ID 62; ORF 010.ng>: g010.pep

```
1 MGFPVRKFDA VIVGGGAGL RAALQLSKSG LNCAVLSKVF PTRSHTVAAQ
51 GGISASLGNV QEDRWDWHMY DTVKGSDWLG DQDAIEFMCR AAPEAVIELE
101 HMGMPFDRVE SGKIYQRPFG GHTAEHGKRA VERACAVADR TGHAMLHTLY
151 QQNVRANTQF FVEWTAQDLI RDENGDVVGV TAMEMETGEV YIFHAKAVMF
201 ATGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG
251 VLITEGVRGE GGILLNADGE RFMERYAPTV KDLASRDVVS RAMAMEIYEG
301 RGCGKNKDHV LLKIDHIGAE KIMEKLPGIR EISIQFAGID PIKDPIPVVP
351 TTHYMMGGIP TNYHGEVVVP QGDEYEVPVK GLYAAGECAC ASVHGANRLG
```

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 63>: m010.seq (PARTIAL)

```
..nTCCAATTAT CCAAATCCGG TCTGAATTGT GCCGTTTTGT CTAAAGTGTT
       CCCGACCCGT TCGCATACCG TAGCGGCGCA GGGCGGTATT TCCGCCTCTn
 51
       TGGGTAATGT GCAGGAAGAC CGTTGGGACT GGCACATGTA CGATACCGTG
      AAAGGTTCCG ACTGGTTGGG CGACCAAGAT GCGATTGAGT TTATGTGCCG
      CGCCGCGCCT GAAGCCGTAA TTGAGTTGGA ACACATGGGT ATGCCTTTTG
201
      ACCGTGTGGA AAGCGGTAAA ATTTATCAGC GTCCTTTCGG CGGCCATACT
251
      GCCGAACACG GTAAACGCGC GGTAGAACGC GYCTGTGCGG TTGCCGACCG
301
      TACAGGTCAT GCGATGCTGC ATACTTTGTA CCAACAAAAC GTCCGTGCCA
351
      ATACGCAATT CTTTGTGGAA TGGACGGCAC AAGATTTGAT TCGTGATGAA
401
451
      AACGGCGATG TCGTCGGCGT AACCGCCATG GAAATGGAAA CCGGCGAAgT
      TTATATTTC CACGCTAAAG CTGTGATGTT TGCTACCGGC GGCGGCGGTC
501
      GTATTTATGC GTCTTCTACC AATGCCTATA TGAATACCGG CGATGGTTTG
551
      GGTATTTGTG CGCGTGCAGG TATCCCGTTG GAAGACATGG AATTCTGGCA
601
      ATTCCAGCCG ACCGGCGTGG CGGGTGCGGG CGTGTTGATT ACCGAA....
```

# This corresponds to the amino acid sequence <SEQ ID 64; ORF 010>: m010.pep (PARTIAL)

```
1 ..XQLSKSGLNC AVLSKVFPTR SHTVAAQGGI SASXGNVQED RWDWHMYDTV
51 KGSDWLGDQD AIEFMCRAAP EAVIELEHMG MPFDRVESGK IYQRPFGGHT
101 AEHGKRAVER XCAVADRTGH AMLHTLYQON VRANTQFFVE WTAQDLIRDE
151 NGDVVGVTAM EMETGEVYIF HAKAVMFATG GGGRIYASST NAYMNTGDGL
201 GICARAGIPL EDMEFWQFQP TGVAGAGVLI TE...
```

### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 65>:

```
a010.seq
         ATGGGCTTTC CTGTTCGCAA GTTTGATGCC GTGATTGTCG GCGGTGGTGG
      51 TGCAGGTTTA CGCGCANCCC TCCAATTATC CAAATCCGGT CTGAATTGTG
    101 CCGTTTTGTC TAAAGTGTTC CCGACCCGTT CGCATACCGT AGCGGCGCAG
    151 GGCGGTATTT CCGCCTCTCT GGGTAATGTG CAGGAAGACC GTTGGGACTG
        GCACATGTAC GATACCGTGA AAGGTTCCGA CTGGTTGGGC GACCAAGATG
         CGATTGAGTT TATGTGCCGC GCCGCGCCTG AAGCCGTAAT TGAGTTGGAA
         CACATGGGTA TGCCTTTTGA CCGTGTGGAA AGCGGTAAAA TTTATCAGCG
     301
    351 TCCTTTCGGC GGCCATACTG CCGAACACGG TAAACGCGCG GTAGAACGCG
    401 CCTGTGCNGT TGCCGACCGT ACAGGTCATG CGATGCTGCA TACTTTGTAC
         CAACAAAATG TCCGTGCCAA TACGCAATTC TTTGTGGAAT GGACGGCACA
        AGATTTGATT CGTGATGAAA ACGGCGATGT CGTCGGCGTA ACCGCCATGG
     551
         AAATGGAAAC CGGCGAAGTT TATATTTTCC ACGCTAAAGC TGTGATGTTT
         GCTACCGGCG GCGGCGGCCG TATTTATGCG TCTTCTACCA ATGCCTATAT
     601
    651 GAATACCGGC GATGGTTTGG GTATTTGTGC GCGTGCAGGT ATCCCGTTGG
    701 AAGACATGGA ATTCTGGCAA TTCCACCCGA CCGGCGTGGC AGGTGCGGGC
    751 GTGTTGATTA CCGAAGGCGT ACGCGGCGAG GGCGGTATTC TGTTGAATGC
        CGACGGCGAA CGCTTTATGG AACGCTATGC GCCGACCGTA AAAGACTTGG
    801
         CTTCTCGCGA CGTTGTTTCC CGCGCGATGG CGATGGAAAT CTACGAAGGT
    901 CGCGGCTGCG GTAAAAACAA AGACCATGTC TTACTGAAAA TCGACCATAT
    951 CGGCGCAGAA AAAATTATGG AAAAACTGCC GGGCATCCGC GAGATTTCCA
    1001 TTCAGTTCGC CGGTATCGAT CCGATTAAAG ACCCGATTCC CGTTGTGCCG
    1051 ACTACCCACT ATATGATGGG CGGTATTCCG ACCAACTACC ATGGCGAAGT
         TGTCGTTCCT CAAGGCGACG AATACGAAGT GCCTGTAAAA GGTCTGTATG
   1151 CGGCAGGTGA GTGCGCCTGT GCTTCCGTAC ACGGTGCGAA CCGCTTGGGT
   1201 ACGAACTCCC TGCTGGACTT AGTGGTATTC GGTAAAGCTG CCGGCGACAG
   1251 CATGATTAAA TTCATCAAAG AGCAAAGCGA CTGGAAACCT TTGCCTGCTA
```

1301 1351 1401 1451 1501 1551 1601 1651 1701 1751	ATTGCACGCC GAGAAGTCAT AAGAGCAAAG CCTAATTGAA AATCACGCGG AACTGGATGA	ACGTTGATGC GGCGTGTTCC GGCGATTGCC TGTGGAATAC GTGGCGAAAG TGCGCACGCT AACATACGCT CACACCAAGC	ATTGCGCCGC GTACTGATGA GAGCGTGTGA CGCGCGTATC CGACTTTGGT TCAGACGACC GTACCATTCA	AGCGTTTGGA GAACTGCAAC GATTCTGAGC AACGTACCGA GAGGCTTTGG GTCTGCCGAA ATCCTGAGCG GATGCCAATA GGAATACATC	AATCAAAGAC AATTGGATAA GCACGTAAAG CGATGATGAA
------------------------------------------------------------------------------	-------------------------------------------------------------------	-----------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------	------------------------------------------------------

# This corresponds to the amino acid sequence <SEQ ID 66; ORF 010.a>:

010.pep			•	•	
1	MGFPVRKFDA	VIVGGGGAGL	RAXLOLSKSG	LNCAVLSKVF	DTD CUMIIA A A
51	GGISASLGNV	QEDRWDWHMY	DTVKGSDWLG	DQDAIEFMCR	DAMAINGATA
<b>1</b> 01	HMGMPFDRVE	SGKIYORPFG	GHTAEHCKRA	VERACAVADR	AAPLAVIELE
151	CONVRANTOF	FVEWTAODLT	POFNEDWICK	TAMEMETGEV	TGHAMLHTLY
201	ATGGGGRTYA	SSTNAVMNTC	DCICTORDAC	TAMEMETGEV	YIFHAKAVMF
251	VITTEGURGE	CCTLINADOR	DGLGTCARAG	IPLEDMEFWQ	FHPTGVAGAG
301	PCCCRNRDUG	JAKEDUTODE.	REMERYAPTV	KDLASRDVVS	RAMAMEIYEG
351	THUMMOOT	LLKIDHIGAE	KIMEKLPGIR	EISIQFAGID	PIKDPIPVVP
	TIMINGGIP	TNYHGEVVVP	OGDEYEVPVK	CLVAACECAC	A CITTLE A STATE OF
401	THOPPDIVAL	GKAAGDSMIK	FIKEOSDWKP	I.DAMACET TO	ODIEDE
451	DGENVDALKK	ELQRSVQLHA	GVFRTDETLS	KCUDEUMATA	EDIMORDET
501	VOVAMITART	EALELDNLIE	VAKATLVSAF	ARKESPCAUA	SUMMEDED
<b>5</b> 51	NWMKHTLYHS	DANTLSYKPV	HTKPLSVEYI	KDVKD11A*	SOURFERDDE
				WE THINK I "	

## m010/a010 98.7% identity over a 231 aa overlap

m010			10	20	30	
m010.pep		XQ:	LSKSGLNCAV	LSKVFPTRSH	TVAAOGGISA	ASXGNV
a010	MGFPVRKFDAVIVGG	GGAGLRAXLQ		111111111	1 1 1 1 1 1 1 1 1 1 1	
	10	20	30	40	50	60
	40 50	60	7.0			
m010.pep	QEDRWDWHMYDTVKG	OU TEANONATE	70 TEMCDA A DE A	80	90	
	11111111111111		HILLILL	ATETEHWGW51	FDRVESGKIY	QRPFG
a010	QEDRWDWHMYDTVKG	DWLGDQDAIL	EFMCRAAPEA	/TELEHMOMDI		
	70	80	90	100	110	120
	100				110	120
m010.pep	100 110	120	130	140	150	
moro.pep	GHTAEHGKRAVERXCA	VADRTGHAMI	HTLYQQNVR	NTQFFVEWTA	QDLIRDENG	DVVGV
a010		111111111	1111111111			
	GHTAEHGKRAVERACA	140	ISO	NTQFFVEWTA 160	QDLIRDENG	
			130	160	170	180
•••	160 170	180	190	200	210	
m010.pep	TAMEMETGEVYIFHAK	AVMFATGGGG	RIYASSTNAY	MNTGDGLGIC	ARAGIPLED	ME FWO
a010			1 1 1 1 1 1 1 1 1 1		11:11:1::::	
4010	TAMEMETGEVYIFHAK 190	AVMFATGGGG 200	RIYASSTNAY	MNTGDGLGIC	ARAGIPLED	MEFWQ
	190	200	210	220	230	240
	220 230					
m010.pep	FQPTGVAGAGVLITE					•
0.4.0	1:11111111111					
a010	FHPTGVAGAGVLITEG	VRGEGGILLN	ADGERFMERY	APTVKDLASR	DVVSRAMAMI	FTVEC
	250	260	270	280	290	300

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae
ORF 010 shows 98.7% identity over a 231 aa overlap with a predicted ORF (ORF 010.ng) from N. gonorrhoeae:

m010.pep/g010.pep

			10	20	30	
m010.pep					TVAAQGGISA	ASXGNV
						1 111
g010	MGFPVRKFDAVIVGG	GGAGLRAALQI	LSKSGLNCAV:	LSKVFPTRSH	TVAAQGGISA	ASLGNV
	10	20	30	40	50	60
	40 50					
	40 50	60	70	80	90	
m010.pep	QEDRWDWHMYDTVKG:	SDWLGDQDAIE	FMCRAAPEA	VIELEHMGMP	FDRVESGKIY	QRPFG
		1111111111111			1111111111	11111
g010	QEDRWDWHMYDTVKG	SDWLGDQDAIE	FMCRAAPEA		FDRVESGKIY	QRPFG
	70	80	90	100	110	120
	100					
44.0	100 110	120	130	140	150	
m010.pep	GHTAEHGKRAVERXC	AVADRTGHAMI	.HTLYQQNVR.	ANTQFFVEWT:	AQDLIRDENG	DVVGV
					1111111111	11111
g010	GHTAEHGKRAVERACA	AVADRTGHAMI	HTLYQQNVR	ANTQFFVEWT:	AQDLIRDENG	DVVGV
	130	140	150	160	170	180
	4.50					
010	160 170	180	190	200	210	
m010.pep	TAMEMETGEVYIFHA	KAVMFATGGGG	RIYASSTNAY	MNTGDGLGI	CARAGIPLED	MEFWQ
01.0					[[[]]]	
g010	TAMEMETGEVYIFHA	CAVMFATGGGG				MEFWQ
	190	200	210	220	230	240
	220 230					
-010						
m010.pep	FOPTGVAGAGVLITE					
~010	:					
g010	FHPTGVAGAGVLITEG	VRGEGGILLN				EIYEG
	250	260	270	280	290	300

### The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 67>: g010-1.seq..

1	ATGGGTTTTC	CTGTTCGCAA	GTTTGATGCC	GTGATTGTCG	GCGGTGGCGG
51	TGCAGGTTTA			CAAATCCGGT	TTGAATTGTG
101	CCGTTTTGTC	TAAAGTGTTC	CCGACCCGCT	CGCATACCGT	AGCGGCGCAG
151	GGCGGTATTT	CCGCCTCTCT	GGGTAATGTG	CAGGAGGACC	GTTGGGACTG
201	GCACATGTAC	GATACCGTGA	AAGGTTCCGA	CTGGCTGGGC	GACCAAGATG
251	CGATTGAGTT	TATGTGTCGC	GCTGCGCCTG	AAGCGGTGAT	TGAGTTGGAA
301	CACATGGGTA	TGCCTTTTGA	CCGCGTTGAA	AGCGGCAAAA	TTTATCAGCG
351	TCCTTTCGGC	GGACATACTG	CCGAACATGG	TAAACGTGCG	GTAGAACGTG
401	CATGTGCGGT	TGCCGACCGT	ACCGGTCATG	CGATGTTGCA	TACTTTGTAC
451	CAACAAAACG	TCCGTGCCAA	TACACAATTC	TTTGTGGAAT	
501	AGATTTGATT	CGTGATGAAA	ACGGCGATGT	CGTCGGCGTA	ACCGCCATGG
551	AAATGGAAAC	GGGCGAAGTT	TATATTTTCC	ACGCCAAGGC	
601	GCTACCGGTG	GCGGCGGTCG	TATTTATGCT	TCTTCTACCA	ATGCTTATAT
651	GAATACCGGT	GACGGTTTGG	GCATTTGCGC	CCGTGCGGGC	
701	AAGATATGGA	ATTCTGGCAA	TTCCACCCGA	CCGGCGTGGC	GGGTGCGGGC
751	GTGTTGATTA	CCGAAGGCGT	ACGCGGCGAG	GGCGGTATTC	TGTTGAACGC
801	CGACGGCGAA	CGCTTTATGG	AACGCTATGC	GCCGACCGTA	
851	CTTCTCGCGA	CGTGGTTTCA	CGCGCGATGG	CGATGGAAAT	CTATGAAGGT
901	CGCGGCTGTG	GTAAAAACAA		TTACTGAAAA	TCGACCATAT
951	CGGTGCAGAA	AAAATTATGG	AAAAACTGCC	GGGCATCCGC	GAGATTTCCA
1001	TTCAGTTTGC	CGGTATCGAT	CCGATTAAAG	ACCCGATTCC	GGTTGTGCCG
1051	ACTACCCACT	ATATGATGGG	CGGCATTCCG	ACCAATTATC	ACGGTGAAGT
1101	TGTTGTTCCG	CAAGGCGACG	AGTACGAAGT	ACCTGTAAAA	GGCCTGTATG
1151	CCGCAGGTGA	GTGCGCCTGT	GCTTCCGTAC	ACGGTGCGAA	CCGTTTGGGT

WO 99/57280

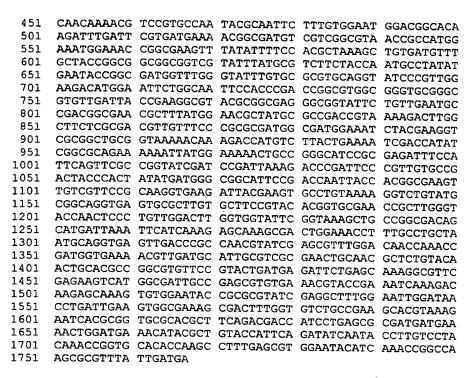
188

1201 ACGAACTCCC TGCTGGACTT GGTGGTGTTC cgcccaaccc cccggtga

```
This corresponds to the amino acid sequence <SEQ ID 68; ORF 010-1.ng>:
        g010-1.pep
               1 MGFPVRKFDA VIVGGGGAGL RAALQLSKSG LNCAVLSKVF PTRSHTVAAQ
                  GGISASLGNV QEDRWDWHMY DTVKGSDWLG DQDAIEFMCR AAPEAVIELE
              51
                  HMGMPFDRVE SGKIYQRPFG GHTAEHGKRA VERACAVADR TGHAMLHTLY
                  QQNVRANTQF FVEWTAQDLI RDENGDVVGV TAMEMETGEV YIFHAKAVMF
             201 ATGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG
             251 VLITEGVRGE GGILLNADGE RFMERYAPTV KDLASRDVVS RAMAMEIYEG
             301 RGCGKNKDHV LLKIDHIGAE KIMEKLPGIR EISIQFAGID PIKDPIPVVP
351 TTHYMMGGIP TNYHGEVVVP OGDEVEVPVK GIVAAGEGAG ASYMGANDIG
                  TTHYMMGGIP TNYHGEVVVP QGDEYEVPVK GLYAAGECAC ASVHGANRLG
             351 TTHYMMGGIP TNYHGE
401 TNSLLDLVVF RPTPR*
 g010-1 / P10444
 sp|P10444|DHSA_ECOLI SUCCINATE DEHYDROGENASE FLAVOPROTEIN SUBUNIT
 gnl|PID|d1015210 (D90711) Succinate dehydrogenase, flavoprotein [Escherichia coli] gi|1786942
 (AE000175) succinate dehydrogenase flavoprotein subunit [Escherichia coli] Length = 588
  Score = 1073 (495.6 bits), Expect = 6.7e-169, Sum P(2) = 6.7e-169
  Identities = 191/303 (63%), Positives = 238/303 (78%)
            1 MGFPVRKFDAVIVXXXXXXXXXXXXXXXSKSGLNCAVLSKVFPTRSHTVAAQGGISASLGNV 60
              M PVR+FDAV++
                                        S+SG CA+LSKVFPTRSHTV+AQGGI+ +LGN
            1 MKLPVREFDAVVIGAGGAGMRAALQISQSGQTCALLSKVFPTRSHTVSAQGGITVALGNT 60
 Sbjct:
           61 QEDRWDWHMYDTVKGSDWLGDQDAIEFMCRAAPEAVIELEHMGMPFDRVESGKIYQRPFG 120
 Query:
               ED W+WHMYDTVKGSD++GDQDAIE+MC+ PEA++ELEHMG+PF R++ G+IYQRPFG
          61 HEDNWEWHMYDTVKGSDYIGDQDAIEYMCKTGPEAILELEHMGLPFSRLDDGRIYQRPFG 120
 Sbjct:
         121 GHTAEHGKRAVERACAVADRTGHAMLHTLYQQNVRANTQFFVEWTAQDLIRDENGDVVGV 180
                         R A ADRTGHA+LHTLYQQN++ +T F EW A DL+++++G VVG
         121 GQSKNFGGEQAARTAAAADRTGHALLHTLYQQNLKNHTTIFSEWYALDLVKNQDGAVVGC 180
Sbjct:
         181 TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ 240
Query:
             TA+ +ETGEV F A+A + ATGG GRIY S+TNA++NTGDG+G+ RAG+P++DME WQ
         181 TALCIETGEVVYFKARATVLATGGAGRIYQSTTNAHINTGDGVGMAIRAGVPVQDMEMWQ 240
Sbict:
         241 FHPTGVAGAGVLITEGVRGEGGILLNADGERFMERYAPTVKDLASRDVVSRAMAMEIYEG 300
Query:
             FHPTG+AGAGVL+TEG RGEGG LLN GERFMERYAP KDLA RDVV+R++ +EI EG
         241 FHPTGIAGAGVLVTEGCRGEGGYLLNKHGERFMERYAPNAKDLAGRDVVARSIMIEIREG 300
Sbjct:
         301 RGC 303
Query:
             RGC
Sbjct:
         301 RGC 303
 Score = 249 (115.0 bits), Expect = 6.7e-169, Sum P(2) = 6.7e-169
 Identities = 53/102 (51%), Positives = 62/102 (60%)
         309 HVLLKIDHIGAEKIMEKLPGIREISIQFAGXXXXXXXXXXXTTHYMMGGIPTNYHGEVV 368
Ouerv:
             H LK+DH+G E + +LPGI E+S FA
                                                       T HYMMGGIPT
         310 HAKLKLDHLGKEVLESRLPGILELSRTFAHVDPVKEPIPVIPTCHYMMGGIPTKVTGQAL 369
Sbjct:
        369 VPQGDEYEVPVKGLYAAGECACASVHGANRLGTNSLLDLVVF 410
                    +V V GL+A GE AC SVHGANRLG NSLLDLVVF
         370 TVNEKGEDVVVPGLFAVGEIACVSVHGANRLGGNSLLDLVVF 411
Sbjct:
```

### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 69>: m010-1.seq..

```
ATGGGTTTTC CTGTTCGCAA GTTTGATGCC GTGATTGTCG GCGGTGGTGG
  1
    TGCAGGTTTA CGCGCAGCCC TCCAATTATC CAAATCCGGT CTGAATTGTG
 51
    CCGTTTTGTC TAAAGTGTTC CCGACCCGTT CGCATACCGT AGCGGCGCAg
101
151 GGCGGTATTT CCGCCTCTCT GGGTAATGTG CAGGAAGACC GTTGGGACTG
    GCACATGTAC GATACCGTGA AAGGTTCCGA CTGGTTGGGC GACCAAGATG
251 CGATTGAGTT TATGTGCCGC GCCGCGCCTG AAGCCGTAAT TGAGTTGGAA
301 CACATGGGTA TGCCTTTTGA CCGTGTGGAA AGCGGTAAAA TTTATCAGCG
351 TCCTTTCGGC GGCCATACTG CCGAACACGG TAAACGCGCG GTAGAACGCG
401 CCTGTGCGGT TGCCGACCGT ACAGGTCATG CGATGCTGCA TACTTTGTAC
```



This corresponds to the amino acid sequence <SEQ ID 70; ORF 010-1>: m010-1.pep..

MGFPVRKFDA VIVGGGGAGL RAALQLSKSG LNCAVLSKVF PTRSHTVAAQ GGISASLGNV QEDRWDWHMY DTVKGSDWLG DQDAIEFMCR AAPEAVIELE HMGMPFDRVE SGKIYQRPFG GHTAEHGKRA VERACAVADR TGHAMLHTLY 101 QQNVRANTQF FVEWTAQDLI RDENGDVVGV TAMEMETGEV YIFHAKAVMF 151 201 ATGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG 251 VLITEGVRGE GGILLNADGE RFMERYAPTV KDLASRDVVS RAMAMEIYEG 301 RGCGKNKDHV LLKIDHIGAE KIMEKLPGIR EISIQFAGID PIKDPIPVVP 351 TTHYMMGGIP TNYHGEVVVP QGEDYEVPVK GLYAAGECAC ASVHGANRLG TNSLLDLVVF GKAAGDSMIK FIKEQSDWKP LPANAGELTR QRIERLDNQT DGENVDALRR ELQRSVQLHA GVFRTDEILS KGVREVMAIA ERVKRTEIKD KSKVWNTARI EALELDNLIE VAKATLVSAE ARKESRGAHA SDDHPERDDE 501 551 NWMKHTLYHS DINTLSYKPV HTKPLSVEYI KPAKRVY*

m010-1 / g010-1 99.5% identity in 410 aa overlap

	10	20	30	40	50	60
m010-1.pep	MGFPVRKFDAVIV	GGGGAGLRAAL	DLSKSGLNCA	VLSKVFPTRS	HTVAAOGGI	SASLENV
	1111111111111			1111111111		111111
g010-1	MGFPVRKFDAVIV	GGGGAGLRAAL	DLSKSGLNCA	VLSKVFPTRS	SHTVAAOGGT	SASTICNY
	10	20	30	40	50	60
	70	80	90	100	110	120
m010-1.pep	QEDRWDWHMYDTV	KGSDWLGDODA	IEFMCRAAPE			TVOPPEC
	- <del>Тининини</del>					
g010-1	QEDRWDWHMYDTV	KGSDWLGDODA	TEFMCRAAPE	AVTELEHMON	(DEUDAEGGE.	TYODDEC
	70	80	90	100	110	120
	-			100	110	120
	130	140	150	160	170	180
m010-1.pep	GHTAEHGKRAVER	ACAVADRTCHA				180
		11111111111		TITITITITI	TACOUTROE	
g010-1	Ghtaehgkraver				·	
	130	140	150			
	230	140	130	160	170	180
	190	200	210	220		
m010-1.pep					230	240
more r.pep	TAMEMETGEVYIF:	HILLIIIIIII	GRIIASSIN	AXMNTGDGLG	ICARAGIPLE	DMEFWQ
g010-1	TAMEMET CENTER	!	700700000		1111111111	111111
A020 T	TAMEMETGEVYIF	UNITED ALL WILD COL	-GRI YASSTN	AYMNTCDGLG	ICARAGIPLE	DMEFWQ

BNSDOCID: <WO 9957280A2 1 >

-

	190	200	210	220	230	240
-010.1	250	260	270	280	290	300
m010-1.pep	FHPTGVAGAGVLIT	EGVRGEGGI:	LLNADGERF <u>M</u> E	RYAPTVKDLA	SRDVVSRAM	AMEIYEG
	111111111111111	11111111	11111111111	1111111111		
g010-1	FHPTGVAGAGVLIT	EGVRGEGGI:	LLNADGERFME	RYAPTVKDL	SRDVVSRAM	METVEC
	250	260	270	280	290	300
			•	200	230	300
	310	320	330	340	250	
m010-1.pep	RGCGKNKDHVLLKI			24U	350	360
			EVELETET	OFACIDALKE	PIPVVPTTHY	MMGGIP
g010-1		11111111		TOTALLI	1111111111	11111
9010 1	RGCGKNKDHVLLKI	DHIGAEKIM		QFAGIDPIKD	PIPVVPTTHY	MMGGIP
	310	320	330	340	350	360
	370	380	390	400	410	420
m010-1.pep	TNYHGEVVVPQGED	YEVPVKGLY	AGECACASVH	GANRLGTNSL	LDLVVFCKAA	CDSMIK
	11111111111111::		1111111111	ELLERIES	11 (11)	0201771/
g010-1	TNYHGEVVVPQGDE	YEVPVKGLY	AGECACASVH	CANRI.CONST.		-
	370	380	390	400		RX.
		500	350	400	410	
	430	440	450			
m010-1.pep			450	460	470	480
moro r.beb	FIKEQSDWKPLPANZ	*GELTKÖKIE	RLDNQTDGEN	<b>VDALRRELO</b> R	Svolhagvfr	TDEILS

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 71>:

```
1 ATGGGCTTTC CTGTTCGCAA GTTTGATGCC GTGATTGTCG GCGGTGGTGG
      TGCAGGTTTA CGCGCANCCC TCCAATTATC CAAATCCGGT CTGAATTGTG
      CCGTTTTGTC TAAAGTGTTC CCGACCCGTT CGCATACCGT AGCGGCGCAG
  101
      GGCGGTATTT CCGCCTCTCT GGGTAATGTG CAGGAAGACC GTTGGGACTG
      GCACATGTAC GATACCGTGA AAGGTTCCGA CTGGTTGGGC GACCAAGATG
  201
  251 CGATTGAGTT TATGTGCCGC GCCGCGCCTG AAGCCGTAAT TGAGTTGGAA
      CACATGGGTA TGCCTTTTGA CCGTGTGGAA AGCGGTAAAA TTTATCAGCG
      TCCTTTCGGC GGCCATACTG CCGAACACGG TAAACGCGCG GTAGAACGCG
 351
      CCTGTGCNGT TGCCGACCGT ACAGGTCATG CGATGCTGCA TACTTTGTAC
      CAACAAAATG TCCGTGCCAA TACGCAATTC TTTGTGGAAT GGACGGCACA
 451
 501 AGATTTGATT CGTGATGAAA ACGGCGATGT CGTCGGCGTA ACCGCCATGG
 551 AAATGGAAAC CGGCGAAGTT TATATTTTCC ACGCTAAAGC TGTGATGTTT
 601 GCTACCGGCG GCGGCGGCCG TATTTATGCG TCTTCTACCA ATGCCTATAT
      GAATACCGGC GATGGTTTGG GTATTTGTGC GCGTGCAGGT ATCCCGTTGG
 651
      AAGACATGGA ATTCTGGCAA TTCCACCCGA CCGGCGTGGC AGGTGCGGGC
 701
 751 GTGTTGATTA CCGAAGGCGT ACGCGGCGAG GGCGGTATTC TGTTGAATGC
      CGACGGCGAA CGCTTTATGG AACGCTATGC GCCGACCGTA AAAGACTTGG
 851 CTTCTCGCGA CGTTGTTTCC CGCGCGATGG CGATGGAAAT CTACGAAGGT
     CGCGGCTGCG GTAAAAACAA AGACCATGTC TTACTGAAAA TCGACCATAT
 901
      CGGCGCAGAA AAAATTATGG AAAAACTGCC GGGCATCCGC GAGATTTCCA
 951
1001
     TTCAGTTCGC CGGTATCGAT CCGATTAAAG ACCCGATTCC CGTTGTGCCG
     ACTACCCACT ATATGATGGG CGGTATTCCG ACCAACTACC ATGGCGAAGT
      TGTCGTTCCT CAAGGCGACG AATACGAAGT GCCTGTAAAA GGTCTGTATG
1101
1151
      CGGCAGGTGA GTGCGCCTGT GCTTCCGTAC ACGGTGCGAA CCGCTTGGGT
1201
      ACGAACTCCC TGCTGGACTT AGTGGTATTC GGTAAAGCTG CCGGCGACAG
      CATGATTAAA TTCATCAAAG AGCAAAGCGA CTGGAAACCT TTGCCTGCTA
1251
1301 ATGCCGGCGA ACTGACCCGC CAACGTATCG AGCGTTTGGA CAATCAAACT
      GATGGTGAAA ACGTTGATGC ATTGCGCCGC GAACTGCAAC GCTCCGTACA
1351
1401
      ATTGCACGCC GGCGTGTTCC GTACTGATGA GATTCTGAGC AAAGGCGTTC
     GAGAAGTCAT GGCGATTGCC GAGCGTGTGA AACGTACCGA AATCAAAGAC
1451
1501 AAGAGCAAAG TGTGGAATAC CGCGCGTATC GAGGCTTTGG AATTGGATAA
1551
     CCTAATTGAA GTGGCGAAAG CGACTTTGGT GTCTGCCGAA GCACGTAAAG
1601 AATCACGCGG TGCGCACGCT TCAGACGACC ATCCTGAGCG CGATGATGAA
1651 AACTGGATGA AACATACGCT GTACCATTCA GATGCCAATA CCTTGTCCTA
     CAAACCGGTG CACACCAAGC CTTTGAGCGT GGAATACATC AAACCGGCCA
1701
1751 AGCGCGTTTA TTGA
```

This corresponds to the amino acid sequence <SEQ ID 72; ORF 010-1.a>: a010-1.pep..

1 MGFPVRKFDA VIVGGGGAGL RAXLQLSKSG LNCAVLSKVF PTRSHTVAAQ 51 GGISASLGNV QEDRWDWHMY DTVKGSDWLG DQDAIEFMCR AAPEAVIELE 101 HMGMPFDRVE SGKIYQRPFG GHTAEHGKRA VERACAVADR TGHAMLHTLY

151 201 251 301 351 401 451 501	QQNVRANTQF ATGGGGRIYA VLITEGVRGE RGCGKNKDHV TTHYMMGGIP TNSLLDLVVF DGENVDALRR KSKVWNTARI NWMKHTLYHS	SSTNAYMNTG GGILLNADGE LLKIDHIGAE TNYHGEVVVP GKAAGDSMIK ELQRSVQLHA EALELDNLIE	DGLGICARAC RFMERYAPTV KIMEKLPGIF QGDEYEVPVF FIKEQSDWKF GVFRTDEILS VAKATLVSAE	F IPLEDME  KDLASRD  EISIQFA  GLYAAGE  LPANAGE  KGVREVM  ARKESRG	FWQ FHPTC VVS RAMAN GID PIKDE CAC ASVHO LTR QRIEF AIA ERVKE AHA SDDHE	EVAGAG MEIYEG PIPVVP SANRLG RLDNQT RTEIKD
m010-1 / a010-1	99.3% ident:	ity in 587 aa	a overlap			
a010-1.pep	10 MGFPVRKFDAVIV	20 /GGGGAGLRAXL(	30 QLSKSGLNCAVL	40 SKVFPTRSHT	50 VAAQGGISA:	60 SLGNV
a010-1	MGFPVRKFDAVIV	/GGGGAGLRAAL( 20	OLSKSGLNCAVL 30	SKVFPTRSHT 40	VAAQGGISA: 50	60
a010-1.pep	70 QEDRWDWHMYDTV	80 /KGSDWLGDQDA	90 EFMCRAAPEAV	100 IELEHMGMPF	110 DRVESGKIYO	120 QRPFG
m010-1	QEDRWDWHMYDTV	/KGSDWLGDQDA1 /KGSDWLGDQDA1		IELEHMGMPF	 DRVESGKIY( 110	 QRPFG 120
a010-1.pep	130 GHTAEHGKRAVEF	140 ACAVADRTGHAN	150 ALHTLYQQNVRA	160 NTQFFVEWTA	170 QDLIRDENGI	180 DVVGV
m010-1	GHTAEHGKRAVER	 PACAVADRTGHAN 140		 NTQFFVEWTA 160	 .QDLIRDENGI 170	  VVGV  180
a010-1.pep	190 TAMEMETGEVYIF	200	210	220	230	240
m010-1	TAMEMETGEVITE		1111111111			ПĨ
	190	200	210	220	230	240
a010-1.pep	250 FHPTGVAGAGVLI	260 TEGVRGEGGILI	270 NADGERFMERY 	280 APTVKDLASR 	290 DVVSRAMAME	300 IYEG
m010-1	FHPTGVAGAGVLI 250	TEGVRGEGGILL 260	NADGERFMERY 270	APTVKDLASR 280	DVVSRAMAME 290	IYEG 300
a010-1.pep	310 RGCGKNKDHVLLK	320 IDHIGAEKIMEK	330 LPGIREISIOF	340 AGIDPIKDPI	350 PVVPTTHYMM	360 IGGIP
m010~1	RGCGKNKDHVLLK	 IDHIGAEKIMEK	 LPGIREISIQF		111111111	1111
	310 370	320 380	330 390	<b>34</b> 0 <b>40</b> 0	350 410	360 420
a010-1.pep	TNYHGEVVVPQGD	EYEVPVKGLYAA	GECACASVHGAL	NRLGTNSLLD	LVVFGKAAGE	SMIK
m010-1	TNYHGEVVVPQGE 370	DYEVPVKGLYAA 380	GECACASVHGAN 390	NRLGTNSLLD	LVVFGKAAGD 410	SMIK 420
a010-1.pep	430 FIKEQSDWKPLPA	440 NAGELTRORIER	450	460 AT.PRET.OPSW	470 OLUNGUERTE	480
m010-1		11111111111	11111111111			1111
	430	440	450	460	470	480
a010-1.pep	490 KGVREVMAIAERV	500 KRTEIKDKSKVW	510 NTARIEALELDN	520 NLIEVAKATL	530 VSAEARKESR	540 GAHA
m010-1				NLIEVAKATL	VSAEARKESR	GAHA
	550	560	570	520 580	530	540
a010-1.pep	SDDHPERDDENWM	KHTLYHSDANTL	SYKPVHTKPLSV	/EYIKPAKRV	YX 	

```
SDDHPERDDENWMKHTLYHSDINTLSYKPVHTKPLSVEYIKPAKRVYX
 m010-1
                       550
                               560
                                                 580
 The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 73>:
       g011.seq
                ATGAAGACAC ACCGCAAGAC CTGCTCTGCG GTGTGTTTTG CTTTTCAGAC
             1
            51
                GGCATCGAAA CCCGCCGTTT CCATCCGACA TCCCAGCGAG GACATCATGA
                GCCTGAAAAC CCGCCTTACC GAAGATATGA AAACCGCGAT GCGCGCCAAA
           101
           151 GATCAAGTTT CCCTCGGCAC CATCCGCCTC ATCAATGCCG CCGTCAAACA
           201 GTTTGAAGTA GACGAACGCA CCGAAGCCGA CGATGCCAAA ATCACCGCCA
           251 TCCTGACCAA AATGGTCAAA CAGCGCAAAG ACGGCGCGAA AATCTACACT
           301 GAAGCCGGCC GTCAGGATTT GGCAGACAAA GAAAACGCCG AAATCGACGT
               GCTGCACCGC TACCTGCCGC AAATGCTCTC CGCCGGCGAA ATCCGCACCG
           401 CCGTCGAAGC AGCCGTTGCC GAAACCGGCG CGGCAGGTAT GGCGGATATG
           451 GGCAAAGTGA TGGTCGTATT GAAAACCCGC CTCGCCGGCA AAGCCGATAT
           501 GGGCGAAGTC AACAAAATCT TGAAAACCGt aCTGACCGCC tga
 This corresponds to the amino acid sequence <SEQ ID 74; ORF 011.ng>:
      g011.pep
                MKTHRKTCSA VCFAFQTASK PAVSIRHPSE DIMSLKTRLT EDMKTAMRAK
                DQVSLGTIRL INAAVKQFEV DERTEADDAK ITAILTKMVK QRKDGAKIYT
            51
                EAGRQDLADK ENAEIDVLHR YLPOMLSAGE IRTAVEAAVA ETGAAGMADM
           101
                GKVMVVLKTR LAGKADMGEV NKILKTVLTA *
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 75>:
      m011.seq (partial)
               ATGAGGACAC ACCGCAAGAC CTGCTCTGCG GTGTGTTTTG CTTTTCAGAC
            51 GGCATCGAAA CCCGCCGTTT CCATCCGACA TCCCAGCGAG GACATCATGA
               GCCTGAAAAT CCGCCTTACC GAAGACATGA AAACCGCGAT GCGCGCCAAA
          151 GACCAAGTTT CCCTCGGCAC CATCCGCCTC ATCAACGCCG CCGTCAAACA
               GTTTGAAGTG GACGAACGCA CCGAAGCCGA CGATGCCAAA ATCACCGCCA
               TCCTGACCAA AATGGTCAAA CAGCGAAAAG ACAGCGCGAA AATCTACACT
          251
          301 GAAGCCGGCC GTCAGGATTT GGCAGACAAA GAAAACGCCG AAATCGAGGT
351 ACTGCACCGC TACCTTCCCC AAATGCTTTC CGCCGGCGAA ATCCGTACCG
          401 AGGTCGAAGC TGCCGTTGCC GAAACCGGCG CGGCAGGTAT GGCGGATATG
          451 GGTAAAGTCA TGGGGCTGCT GAAAACCCGC CTCGCAGGTA AAGCCGA...
This corresponds to the amino acid sequence <SEQ ID 76; ORF 011>:
     m011.pep (partial)
            1 MRTHRKTCSA VCFAFQTASK PAVSIRHPSE DIMSLKIRLT EDMKTAMRAK
           51 DQVSLGTIRL INAAVKQFEV DERTEADDAK ITAILTKMVK QRKDSAKIYT
          101 EAGRQDLADK ENAEIEVLHR YLPOMLSAGE IRTEVEAAVA ETGAAGMADM
               GKVMGLLKTR LAGKA....
Computer analysis of this amino acid sequence gave the following results:
Homology with a predicted ORF from N. gonorrhoeae
ORF 011 shows 95.8% identity over a 165 aa overlap with a predicted ORF (ORF 011.ng)
from N. gonorrhoeae:
     m011/q011
                                    20
                                             30
                 MRTHRKTCSAVCFAFQTASKPAVSIRHPSEDIMSLKIRLTEDMKTAMRAKDQVSLGTIRL
     m011.pep
                  MKTHRKTCSAVCFAFQTASKPAVSIRHPSEDIMSLKTRLTEDMKTAMRAKDQVSLGTIRL
     g011
                          10
                                             30
                                                       40
                          70
                                    80
                                             90
                                                      100
                                                                110
                 INAAVKQFEVDERTEADDAKITAILTKMVKQRKDSAKIYTEAGRQDLADKENAEIEVLHR
    m011.pep
                 INAAVKQFEVDERTEADDAKITAILTKMVKQRKDGAKIYTEAGRQDLADKENAEIDVLHR
    g011
```

	70	80	90	100	110	120
	130	140	150	160		
m011.pep	YLPQMLSAGEIRT	EVEAAVAETG	AAGMADMGKVM	GLLKTRLAG	A)	
		3		:	l	
g011	YLPQMLSAGEIRT	AVEAAVAETG	AAGMADMGKVM	VVLKTRLAGI	CADMGEVNKII	LKTVLTA
_	130	140	150	160	170	180
g011	x					

## The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 77>: g012.seq

```
ATGCTCGCCC GTCGCTATTT TTTCAATATC CAACCCGGGG CGGTTTTCAC
  1
51
     TGACAAACTG CTTGAACAAC TGATGCGTTT CCTCCAGTTC CTGCCGGAAT
101
    TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCGCTT
151
    AAATTCGCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTTCAACa
    gGcggTGGAT ATTCGgcact tccgCcacca cacccaccga accgatgacc
    gcaaacggaG CGGAAACAAT TTTATCCGCc acacacgcca tcatatagcc
251
    gcCGCTTGCC GCGACCTTAT CGAcggcgac ggTCAGCGGA ATATTGCGTT
301
    CGCGCAAACG CCTAAGCTGC GAAGCCGCCA AACCGTAACC GTGAACCACG
401
    CCGCCCGGAC TTTCCAATCT GAGCAGAACC TCATCTTCAG GCTTGGCAAT
    CAAAAGCACC GCCGTAATCT CATGACGCAA GGATTCTACG GCGTGTGCAT
451
    ACAAATCGCC GTCAAAATCC AACACAAAAA GGCGGGATTT TTGCGTTTCG
501
    GCAGATTTCT CCCCGCCCTC CTTCAAACGC TTTTTCTCTG CTTTGGCTTC
551
    CGCCTTTTCC TTTTTCTTTT CTTTTTTTC CTGATGTTTT GTCTCTTCCT
651
    CGCTTAA
```

## This corresponds to the amino acid sequence <SEQ ID 78; ORF 012.ng>: g012.pep

1 MLARRYFFNI QPGAVFTDKL LEQLMRFLQF LPEFLFALFR IFTHKSNRAL 51 KFARRHHIHI NIMFFQQAVD IRHFRHHTHR TDDRKRSGNN FIRHTRHHIA 101 AACRDLIDGD GQRNIAFAQT PKLRSRQTVT VNHAARTFQS EQNLIFRLGN 151 QKHRRNLMTQ GFYGVCIQIA VKIQHKKAGF LRFGRFLPAL LQTLFLCFGF 201 RLFLFLFFFF LMFCLFLA*

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 79>: m012.seq

ATGCTCGCCC GTTGCCACTT CCTCAATATC CAATTGAGGG CGGTTCTCGC 51 TGACAAACTG CTTGAACAAC TGATGCGTTT CCTCCAGTTC CTGTCGGAAT 101 TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCGCTT AAATTCGCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTTCAACA 151 201 GGCGGTGGAT ATTCGGTACT TCCGCCACCA CACCCACCGA ACCGACAATC 251 GCAAACGGAG CGGAAGCAAT TTTATCCGCC ACACACGCCA TCATATAACC 401 451 501 nnnnnnnn nnnnnnnn AACACAAAAA GGCGTGATTT nTGCGTTTCG GCAGATTTCT CCCCACCCTC CTTCAAACGT TTTTCcTCTG CTTTGGCTTC 551 601 CGCCTTTCC TTTTCTTTT CCTCTTTTC CTGATGTTGT GCCTCTTCCC 651 CGCTTAA

## This corresponds to the amino acid sequence <SEQ ID 80; ORF 012>: m012.pep

```
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 81>:
       a012.seg
                ATGCTCGCCC GTTGCCACTT CCTCAATATC CAATTGAGGG CGGTTCTCGC
             1
                TGACAAACTG CTTGAACAAC TGATGCGTTT CCTCCAGTTC CTGTCGGAAT
            51
                TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCGCTT
            101
               AAATTCGCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTTCAACA
           151
               GGCGGTGGAT ATTCGGTACT TCCGCTACAA CACCCACCGA ACCGACAATC
                GCAAACGGAG CGGAAACAAT TTTATCCGCC ACACACGCCA TCATATAACC
               ACCGCTCGCC GCCACCTTAT CGACGGCGAC GGTCAGCGGA ATATTGCGTT
           301
                CGCGCAAACG CCTAAGCTGC GAAGCCGCCA AACCGTAACC GTGAACCACG
                CCGCCCGGAC TTTCCAATCT AAGCAGAACC TCATCTTCAG GCTTGGCAAT
           401
                CAAAAGCACC GCCGTAATCT CATGACGCAA GGATTCTACG GCGTGTGCAT
           451
           501 ACAAATCGCC GTCAAAATCC AACACAAAAA GGCGGGATTT TTGCGTTTCG
                GAAGATTTCT CCCCACCCTC CTTCAAACGC TTTTTCTCTG CTTTGGCTTC
                CGCCTTTTCC TTTTTCTTTT CCTCTTTTTC CTGATGTTTT GCCTCTTCCC
           651
                CGCTTAA
 This corresponds to the amino acid sequence <SEQ ID 82; ORF 012.a>:
      a012.pep
               MLARCHFLNI QLRAVLADKL LEQLMRFLQF LSEFLFALFR IFTHKSNRAL
             7
               KFARRHHIHI NIMFFQQAVD IRYFRYNTHR TDNRKRSGNN FIRHTRHHIT
            51
               TARRHLIDGD GORNIAFAOT PKLRSROTVT VNHAARTFOS KONLIFRLGN
           101
               QKHRRNLMTQ GFYGVCIQIA VKIQHKKAGF LRFGRFLPTL LQTLFLCFGF
               RLFLFLFFF LMFCLFPA*
             64.2% identity over a 218 aa overlap
 m012/a012
                         10
                                   20
                                            30
                                                     40
                  MLARCHFLNIQLRAVLADKLLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI
                                                                        60
      m012.pep
                  MLARCHFLNIQLRAVLADKLLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI
      a012
                         10
                                            30
                                                     40
                                                               50
                                                                        60
                                   80
                                            90
                                                              110
                  NIMFFQQAVDIRYFRHHTHRTDNRKRSGSNFIRHTRHHITAARXXXXXXXXXXXXXXXXXX
     m012.pep
                  NIMFFQQAVDIRYFRYNTHRTDNRKRSGNNFIRHTRHHITTARRHLIDGDGQRNIAFAQT
     a012
                         70
                                  80
                                            90
                                                    100
                                                              110
                                                                       120
                        130
                                 140
                                           150
                                                    160
     m012.pep
                 180
                 PKLRSRQTVTVNHAARTFQSKQNLIFRLGNQKHRRNLMTQGFYGVCIQIAVKIQHKKAGF
     a012
                        130
                                 140
                                           150
                                                    160
                                                             170
                        190
                                 200
                                           210
                                                   219
     m012.pep
                 XRFGRFLPTLLQTFFLCFGFRLFLFLFLFLMLCLFPAX
                  a012
                 LRFGRFLPTLLQTLFLCFGFRLFLFLFLFFLMFCLFPAX
                        190
                                 200
                                           210
Computer analysis of this amino acid sequence gave the following results:
Homology with a predicted ORF from N. gonorrhoeae
ORF 012 shows 58.7% identity over a 218 aa overlap with a predicted ORF (ORF 012.ng)
from N. gonorrhoeae:
```

30

MLARCHFLNIQLRAVLADKLLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI

MLARRYFFNIQPGAVFTDKLLEQLMRFLQFLPEFLFALFRIFTHKSNRALKFARRHHIHI

40

50

m012/g012

m012.pep

g012

10

	:	10 :	20 3	0 40	50	60
		70 I	30 9	0 100	110	120
m012.pep	NIMFFQQAV	VDIRYFRHHTI	HRTDNRKRSGS	NFIRHTRHHIT	AARXXXXXXXX	
g012				: NFIRHTRHHIA	II AACRDLIDGDGQ	: RNIAFAOT
-			30 9		110	120
			10 15		170	180
m012.pep	XXXXXXXX :	XXXXXXXXXX : :	(XXXXXXXXXX	: xxxxxxxxx :	XXXXXXXXXXX	XANKKAXF
g012					GFYGVCIQIAVK	IOHKKAGF
	13	30 14	10 15	0 160	170	180
m012.pep			00 21 GFRLFLFLFLF			
	1111111:	:1111:1111.	111111111111111111111111111111111111111	111:111		
g012			FRLFLFLFFF 00 21			
The following p		quence was	identified in	n <i>N. meningi</i>	tidis <seq i<="" td=""><td>D 83&gt;:</td></seq>	D 83>:
m012-1.se	eq ATGCTCGCCC G	TTTGCCACTT	ССТСАВТАТС	СААТТСАССС	СССТТС <b>Т</b> ССС	
51	TGACAAACTG	CTTGAACAAC	TGATGCGTTT	CCTCCAGTTC	CTGTCGGAAT	
101	TTCTGTTTGC C					
151 201	AAATTCGCCC G					
201 251	GGCGGTGGAT A					
301	GCCGCTCGCC G					
351	CGCGCAAACG C					
401	CCGCCCGGAC T					
451	CAAAAGCACC G	GCCGTAATCT	CATGACGCAA	GGATTCTACG	GCGTGTGCAT	
501	ACAAATCGCC C					
551	GCAGATTTCT C					
601	CGCCTTTTCC I	PTTTTCTTTT	CCTCTTTTTC	CTGATGTTTT	GCCTCTTCCC	
651	CGCTTAA					
This correspond		acid seque	nce <seq ii<="" td=""><td>D 84; ORF 0</td><td>12-1&gt;:</td><td></td></seq>	D 84; ORF 0	12-1>:	
m012-1.pe						
. 1 51	MLARCHFLNI Ç KFARRHHIHI N					
101	AARRHLIDGD G					
	QKHRRNLMTQ G					
201	RLFLFLFLFF I				26.2120101	
<b>m012-1/g012</b> 9	1.7% identity	in 218 aa o	verlap			
			20 30		50	60
m012-1.pe	_				FTHKSNRALKF	
g012	: :  MI DDVEEN					
g012			20 3(		50	60
m012-1.pe			30 91 ARTONRKRSCSI		110 AARRHLIDGDGQ	120 DNIA EAOT
m012-1.pe					AKKHTIDGDGÖ	
g012	NIMFFOOAV	/DIRHFRHHT	IRTDDRKRSGNI	NFIRHTRHHIA	AACRDLIDGDGQ	RNIAFAOT
<b>-</b>			30 90		110	120
	13	30 14	10 150	160	170	180
m012-1.pe					FYGVCIQIAVK	
-	1111111	[]]]]		1111111111	111111111111	1111111
g012	PKLRSRQTV	/TVNHAARTF(	QSEQNLIFRLGI	NQKHRRNLMTQ(	FYGVCIQIAVK	IQHKKAGF

```
130
                                  140
                                           150
                                                    160
                                                             170
                                                                       180
                         190
                                  200
                                           210
                                                   219
                  LRFGRFLPTLLQTLFLCFGFRLFLFLFLFFLMFCLFPAX
      m012-1.pep
                   LRFGRFLPALLQTLFLCFGFRLFLFLFFFFLMFCLFLAX
      a012
                         190
                                           210
 The following partial DNA sequence was identified in N. meningitidis <SEQ ID 85>:
 a012-1.seq
               ATGCTCGCCC GTTGCCACTT CCTCAATATC CAATTGAGGG CGGTTCTCGC
               TGACAAACTG CTTGAACAAC TGATGCGTTT CCTCCAGTTC CTGTCGGAAT
            51
               TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCGCTT
           101
               AAATTCGCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTTCAACA
           151
               GGCGGTGGAT ATTCGGTACT TCCGCTACAA CACCCACCGA ACCGACAATC
           201
          251
               GCAAACGGAG CGGAAACAAT TTTATCCGCC ACACACGCCA TCATATAACC
          301 ACCGCTCGCC GCCACCTTAT CGACGCGAC GGTCAGCGGA ATATTGCGTT
               CGCGCAAACG CCTAAGCTGC GAAGCCGCCA AACCGTAACC GTGAACCACG
          351
               CCGCCCGGAC TTTCCAATCT AAGCAGAACC TCATCTTCAG GCTTGGCAAT
          401
               CAAAAGCACC GCCGTAATCT CATGACGCAA GGATTCTACG GCGTGTGCAT
          451
               ACAAATCGCC GTCAAAATCC AACACAAAAA GGCGGGATTT TTGCGTTTCG
          501
               GAAGATTTCT CCCCACCCTC CTTCAAACGC TTTTTCTCTG CTTTGGCTTC
               CGCCTTTTCC TTTTTCTTTT CCTCTTTTTC CTGATGTTTT GCCTCTTCCC
               CGCTTAA
 This corresponds to the amino acid sequence <SEQ ID 86; ORF 012-1.a>:
     a012-1.pep
              MLARCHFLNI QLRAVLADKL LEQLMRFLQF LSEFLFALFR IFTHKSNRAL
            1
           51
               KFARRHHIHI NIMFFQQAVD IRYFRYNTHR TDNRKRSGNN FIRHTRHHIT
          101
              TARRHLIDGD GQRNIAFAQT PKLRSRQTVT VNHAARTFQS KQNLIFRLGN
               QKHRRNLMTQ GFYGVCIQIA VKIQHKKAGF LRFGRFLPTL LQTLFLCFGF
          201
              RLFLFLFF LMFCLFPA*
a012-1/m012-1
              97.2% identity in 218 aa overlap
                                           30
                 MLARCHFLNIQLRAVLADKLLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI
     a012-1.pep
                 MLARCHFLNIQLRAVLADKLLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI
     m012-1
                                  20
                                           30
                                                    40
                        70
                                           90
                                                   100
                                                            110
                 NIMFFQQAVDIRYFRYNTHRTDNRKRSGNNFIRHTRHHITTARRHLIDGDGQRNIAFAQT
     a012-1.pep
                 NIMFFQQAVDIRYFRHHTHRTDNRKRSGSNFIRHTRHHITAARRHLIDGDGQRNIAFAQT
     m012-1
                        70
                                 80
                                           90
                                                   100
                                                            110
                       130
                                 140
                                          150
                                                   160
                                                            170
                 PKLRSRQTVTVNHAARTFQSKQNLIFRLGNQKHRRNLMTQGFYGVCIQIAVKIQHKKAGF
     a012-1.pep
                  XKLRSRQTVTVNHAARTFQSEQNLIFRLGNQKHRRNLMTQGFYGVCIQIAVKIQHKKAGF
     m012-1
                       130
                                140
                                         150
                                                   160
                                                            170
                                                                     180
                                200
                                         210
                 LRFGRFLPTLLQTLFLCFGFRLFLFLFLFFLMFCLFPAX
     a012-1.pep
                 m012-1
                 LRFGRFLPTLLQTLFLCFGFRLFLFLFLFFLMFCLFPAX
                       190
                                200
The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 87>:
    g013.seq
```

aTgcctttga ccatgctgtg cagcaGGAcg tGCGGTTtgt tcataataca

gtCcgaccGG AAAagcggAG GAAaCGCAGT GCCGCGCCCT TCCCCTTTCT

TGCCGTGGCA GGCGATGCag tTgGATTCGT ACACTTTTTG CCCTTTtGtc

1

51

```
151 atgatGCTgt tgtcggCGGC AGAAGCgGCG GcgCAGAGGC AGCACAAGAT
          201 GAAGGCGGTC GGCAGTCGGG TTGTGTtcat tGgcgTTTCC cctaatgttt
               tgaaaccttg ttttttgatt Ttgcctttac ggggtgaaaa gttttTtgg
          301 cccaaatccg gaatttag
This corresponds to the amino acid sequence <SEQ ID 88; ORF 013.ng:
     g013.pep
               MPLTMLCSRT CGLFIIQSDR KSGGNAVPRP SPFLPWQAMQ LDSYTFCPFV
            1
           51
               MMLLSAAEAA AQRQHKMKAV GSRVVFIGVS PNVLKPCFLI LPLRGEKFFW
          101
               PKSGI*
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 89>:
            1
               ATGCCTTTGA CCATGCTGTG CAGCAGCACC TGCGGTTTTT TCATGATGAA
           51
              GTCGGAGCGG TAGAGCGGCG GAAACATGGT TCCGCGGCCT TCGCCCTTTT
          101
               TGCCGTGGCA GGCGACGCAG TTGGATTCGT ACACTTTTTG CCCTTTTGTC
               ATGATGCTGT TGTCGGCGGC AGAAGCGGCG GCGCAGAAGC AGCCCAAGAC
          151
               GAGGGCGGTC GGCAGTCGGG TTGTGTTCAT TGGTGTTTCC TTCATGTTTG
          201
              AAACCTTGTT GTTGATTTTG CGTAGCGGGT GAAAGATTTT TTTGCCGAAT
          251
          301
              CAGTAG
This corresponds to the amino acid sequence <SEQ ID 90; ORF 013>:
     m013.pep
               MPLTMLCSST CGFFMMKSER XSGGNMVPRP SPFLPWQATQ LDSYTFCPFV
              MMLLSAAEAA AQKQPKTRAV GSRVVFIGVS FMFETLLLIL RSGXKIFLPN
           51
          101
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 91>:
     a013.seq
               ATGCCTTTGA CCATGCTGTG CAGCAGCACC TGCGGTTTTT TCATGATGAA
           1
           51
              GTCGGAGCGG TAGAGCGGCG GAAACATGGT TCCGCGGCCT TCGCCCTTTT
          101
              TGCCGTGGCA GGCGACGCAG TTGGATTCGT ACACTTTTTG CCCTTTTGTC
          151
              ATGATGCTGT TGTCGGCGGC AGAAGCGGCG GCGCAGAGGC AGCCCAAGAC
          201
              GAGGGCGGTC GGCAGTCGGG TTGTGTTCAT TGGTGTTTCC TTAATGTTTG
          251 AAACCTTGTT GTTGATTTTG CGTAGCGGGT GAAAGATTTT CTTGCCGAAT
          301
              CGGTAG
This corresponds to the amino acid sequence <SEQ ID 92; ORF 013.a>:
              MPLTMLCSST CGFFMMKSER *SGGNMVPRP SPFLPWQATQ LDSYTFCPFV
           51
              MMLLSAAEAA AQRQPKTRAV GSRVVFIGVS LMFETLLLIL RSG*KIFLPN
          101
              R*
            97.0% identity over a 101 aa overlap
m013/a013
                                   20
                                             30
                                                       40
                                                                50
     m013.pep
                 MPLTMLCSSTCGFFMMKSERXSGGNMVPRPSPFLPWQATQLDSYTFCPFVMMLLSAAEAA
                  {\tt MPLTMLCSSTCGFFMMKSERXSGGNMVPRPSPFLPWQATQLDSYTFCPFVMMLLSAAEAA}
     a013
                         10
                                                       40
                                                                50
                                                                          60
                                   80
                                             90
     m013.pep
                 AQKQPKTRAVGSRVVFIGVSFMFETLLLILRSGXKIFLPNQX
                  a013
                 AQRQPKTRAVGSRVVFIGVSLMFETLLLILRSGXKIFLPNRX
                         70
                                   80
                                             90
```

Computer analysis of this amino acid sequence gave the following results: Homology with a predicted ORF from N. gonorrhoeae

ORF 013 shows 73.3% identity over a 101 aa overlap with a predicted ORF (ORF 013.ng) from N. gonorrhoeae:

BNSDOCID: <WO__9957280A2 __>

```
m013/g013
```

```
20
                               30
                                      40
         MPLTMLCSSTCGFFMMKSERXSGGNMVPRPSPFLPWQATQLDSYTFCPFVMMLLSAAEAA
m013.pep
          MPLTMLCSRTCGLFI1QSDRKSGGNAVPRPSPFLPWQAMQLDSYTFCPFVMMLLSAAEAA
g013
                       20
                               30
                                      40
                70
                       80
                                90
                                       100
         AQKQPKTRAVGSRVVFIGVSF-MFETLLLILR-SGXKIFLPNQX
m013.pep
         q013
         AQRQHKMKAVGSRVVFIGVSPNVLKPCFLILPLRGEKFFWPKSGIX
               70
                       80
                               90
                                     100
```

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 93>: g015.seg

```
1 ATGCAGTATC TGATTGTCAA ATACAGCCAT CAAATCTTCG TTACCATCAC
 51 CATTTTGGTA TTCAACATCC GTTTTTTCCT ACTTTGGAAA AATCCAGAAA
101 AGCCCTTGGT CGGCTTTTGG AAAGCACTGC CCCACCTCAA CGACACGATG
151 CTGCTGTTTA CGGGATTGTG GCTGATGAAG ATTACCCATT TCTCCCCGTT
201 CAACGCGCCT TGGCTCGGCA CAAAAATCCT GCTCCTGTTC GCCTACATCG
    CACTGGGCAT GGTAATGATG CGCGCCCGTC CGCGTTCGAC CAAGTTCTAC
    ACCGTTTACC TGCTCGCTAT GTGTTGCATC GCCTGCATCG TTTACCTTGC
351 CAAAACCAAA GTCCTGCCAT TCTGA
```

### This corresponds to the amino acid sequence <SEQ ID 94; ORF 015.ng>: g015.pep

- 1 MQYLIVKYSH QIFVTITILV FNIRFFLLWK NPEKPLVGFW KALPHLNDTM
  - 51 LLFTGLWLMK ITHFSPFNAP WLGTKILLLF AYIALGMVMM RARPRSTKFY 101 TVYLLAMCCI ACIVYLAKTK VLPF*

## The following partial DNA sequence was identified in N. meningitidis <SEQ ID 95>:

```
m015.seq
           (partial)
          . . AAAATCAGAA AAGCCTTGGC GGGCTTTTGG AAGGCACTGC CCCACCTTAA
            CGACACCATG CTGCTGTTTA CGGGATTGTG GCTGATGAAA ATTACCCATT
            TCTCCCCGTT CAACGCGCCT TGGCTCGGTA CAAAAATCCT GCTTCTGCTC
     101
            GCCTATATCG CATTGGGTAT GATGATGATG CGCGCCCGTC CGCGTTCGAC
     151
     201
            CAAGTTCTAC ACCGTTTACC TGCTCGCCAT GTGTTGCGTC GCCTGCATCG
     251
            TTTACCTTGC CAAAACCAAA GTCCTGCCTT TCTGA
```

## This corresponds to the amino acid sequence <SEQ ID 96; ORF 015:

m015.pep (partial)

- 1 ..KIRKALAGFW KALPHLNDTM LLFTGLWLMK ITHFSPFNAP WLGTKILLLL
- AYIALGMMMM RARPRSTKFY TVYLLAMCCV ACIVYLAKTK VLPF*

### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 97>: a015.seq

```
ATGCAGTATC TGATTGTCAA ATACAGCCAT CAAATCTTCG TTACCATCAC
  1
     CATTTTGGTA TTCAACATCC GTGTTTTCNT ACTTTGGAAA AATCCAGAAA
 51
101 AGCCCTTGGC GGGCTTTTGG AAGGCACTGC CCCACCTTAA CGACACCATG
151 CTGCTGTTTA CGGGATTGTG GCTGATGAAA ATTACCCATT TCTCCCCGTT
201 CAACGCGCCT TGGCTCGGTA CAAAAATCCT GCTTCTGCTC GCCTATATCG
251 CATTGGGTAT GATGATGATG CGCGCCCGTC CGCGTTCGAC CAAGTTCTAC
301 ACCGTTTACC TGCTCGCCAT GTGTTGCCTC ACCGTCATCAC
     ACCGTTTACC TGCTCGCCAT GTGTTGCCTC ACCTGCATCG TTTACCTTGC
351 CAAAACCAAA GTCCTGCCTT TCTGA
```

This corresponds to the amino acid sequence <SEQ ID 98; ORF 015.a>: a015.pep

عاملك كالمنا المميين الممال الأجالي

- 1 MQYLIVKYSH QIFVTITILV FNIRVFXLWK NPEKPLAGFW KALPHLNDTM
- 51 LLFTGLWLMK ITHFSPFNAP WLGTKILLLL AYIALGMMMM RARPRSTKFY
  101 TVYLLAMCCL TCIVYLAKTK VLPF*

96.7% identity over a 91 aa overlap m015/a015

				10	20	30
m015.pep			KIRK	ALAGFWKALPH	LNDTMLLFT	GLWLMKITH
				11111111		
a015	LIVKYSHQIFVI				LNDTMLLFT	GLWLMKITH
	10	20	30	40	50	60
	40	50	60	70	80	90
m015.pep	FSPFNAPWLGTK	CILLLLAYIAI	GMMMMRAR P	RSTKFYTVYLL	AMCCVACIV	YLAKTKVLP
		1111111111	11111111	11111111111	1111::111	11111111
a015	FSPFNAPWLGTK	CILLLLAYIAI	GMMMMRARP	RSTKFYTVYLI	AMCCLTCIV	YLAKTKVLP
	70	80	90	100	110	120
m015.pep	FX					
	1.1					
a015	FX					

Computer analysis of this amino acid sequence gave the following results:

### Homology with a predicted ORF from N. gonorrhoeae

ORF 015 shows 94.5% identity over a 91 aa overlap with a predicted ORF (ORF 015.ng) from N. gonorrhoeae:

m015/g015

				10	20	30
m015.pep			KIRK	ALAGFWKALP	HLNDTMLLFT	GLWLMKITH
			!			
g015	LIVKYSHQIFVT	ITILVFNIR	FFLLWKNPEK	PLVGFWKALP:	HLNDTMLLFT	GLWLMKITH
	10	20	30	40	50	60
	40	50	60	70	80	90
m015.pep	FSPFNAPWLGTK	ILLLLAYIA	LGMMMMRARP	RSTKFYTVYL		
• •		1111:111		111111111		
g015	FSPFNAPWLGTK	ILLLFAYIA	LGMVMMRARP	RSTKFYTVYL	LAMCCIACIV	YLAKTKVLP
	70	80	90	100	110	120
m015.pep	FX					
,	TI *					
g015	FX					

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 99>: g018.seq

1	atGCAGCAGG	GGCagttggt	tggacgcgtc	gcccgcaata	AAGATATGCG
51	GAATgctggt	CTGCATggtC	AGCGGATCGG	CAACGGGtac	gccgcgcgcg
101	tctttgTCGA	TATTGATGTT	TTCCAAACCG	ATATtgTCAA	CGTTCGGACG
151			TATATTCGGC		
201	CATCCTGCTC	CCAATGGACT	tctACATTGC	CGTCTGCGTC	GAGTTTGACC
251	TCGGTTTTAG	CATCCAGATG	CAGTTTCAAT	tctTCTCCGA	ACACGGCTTT
301	CGCCTCGTCT	GA			

This corresponds to the amino acid sequence <SEQ ID 100; ORF 018.ng>: g018.pep

- 1 MQQGQLVGRV ARNKDMRNAG LHGQRIGNGY AARVFVDIDV FQTDIVNVRT
- 51 ATYGCQHIFG NKYAFFAILL PMDFYIAVCV EFDLGFSIQM QFQFFSEHGF

101 RLV*

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 101>:

1 ATGCAGCAGA GGCAGTTGGT TGGACGCATC GCCTGCGATG AAGATATGCG
51 GAATACTGGT CTGCATGGTC AGCGGGTCGG CAACAGGTAC GCCGCGCGCA
101 TCTTTTTCGA TATTGATATT TTCCAAACCG ATATTGTCAA CGTTCGGACG
151 GCGGCCCACG GCTGCCAGCA TATATTCGGC AACAAATACG CCTTTTTCGC
201 CATCCTGCTC CCAATGGACT TCTACATTGC CGTCTGCATC GAGTTTGACC
251 TCGGTTTTAG CATCCAGATG CAGTTTCAAT TCTTCGCCGA ACACGGCGTT
301 CGCCTCGTCT GA

# This corresponds to the amino acid sequence <SEQ ID 102; ORF 018>: m018.pep

1 MQQRQLVGRI ACDEDMRNTG LHGQRVGNRY AARIFFDIDI FQTDIVNVRT 51 AAHGCQHIFG NKYAFFAILL PMDFYIAVCI EFDLGFSIQM QFQFFAEHGV 101 RLV*

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 103>:

1 ATGCAGCAGG GGCAGTTGGT TGGACGCGTC GCCCGCAATA AAGATATGCG
51 GAATACTGGT CTGCATAGTC AGCGGATCGG CAACGGGTAC GCCGCGCGCA
101 TCTTTTTCGA TATTGATGTT TTCCAAACCG ATATTGTCAA CGTTCGGACG
151 GCGCCTACG GCTGCCAGCA TATATTCGCC AACAAATACG CCTTTTTCGC
201 CATCCTGCTC CCAATGGACT TCTACATTGC CGTCTGCGTC GAGTTTGGCC
251 TCGGTTTTAG CATCCAAATG CAGTTTCAAT TCTTCACCGA ACACGGCTTT
301 CGCCTCGTCT GA

## This corresponds to the amino acid sequence <SEQ ID 104; ORF 018.a>:

1 MQQGQLVGRV ARNKDMRNTG LHSQRIGNGY AARIFFDIDV FQTDIVNVRT 51 AAYGCQHIFG NKYAFFAILL PMDFYIAVCV EFGLGFSIQM QFQFFTEHGF 101 RLV*

### m018/a018 86.4% identity over a 103 aa overlap

20 30 40 50  ${\tt MQQRQLVGRIACDEDMRNTGLHGQRVGNRYAARIFFDIDIFQTDIVNVRTAAHGCQHIFG}$ m018.pep MQQGQLVGRVARNKDMRNTGLHSQRIGNGYAARIFFDIDVFQTDIVNVRTAAYGCQHIFG a018 10 20 30 40 70 80 90 100 m018.pep NKYAFFAILLPMDFYIAVCIEFDLGFSIQMQFQFFAEHGVRLVX a018 NKYAFFAILLPMDFYIAVCVEFGLGFSIQMQFQFFTEHGFRLVX 70 80 90 100

# Computer analysis of this amino acid sequence gave the following results: Homology with a predicted ORF from N. gonorrhoeae

ORF 018 shows 84.5% identity over a 103 aa overlap with a predicted ORF (ORF 018.ng) from N. gonorrhoeae:

m018/g018

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 105>:
q019.seq (partial)

```
1 ..etgetggegg ceetggtget tgeeggtgt tettegACAA ACAcacTGCC
51 AGCCGGCAAG ACCCCGGCAG ACAATATAGA AActgeCgAC CTTTCGGCAA
101 GCGTTCCCAC cegeCCTGCC GAACCGGAAG GAAAAACGCT GGCAGATTAC
151 GGCGGCTACC CGTCCGCACT GGATGCAGTG AAACAGAACA ACGATGCGGC
201 AGCCGCCGCC TATTTGGAAA Acgcaggaga cagCGcgatg gcGGAAAatg
```

251 tccgcaagga gtgGCTGa

This corresponds to the amino acid sequence <SEQ ID 106; ORF 019.ng>:

g019.pep (partial)

1 ..LLAALVLAAC SSTNTLPAGK TPADNIETAD LSASVPTRPA EPEGKTLADY

51 GGYPSALDAV KQNNDAAAAA YLENAGDSAM AENVRKEWL*

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 107>:

```
m019.seq (partial)
         ATGTACCTAC CCTCTATGAA GCATTCCCTG CCGCTGCTGG CGGCCCTGGT
      1
      51
         GCTTGCCGCG TGTTCTTCGA CAAACACACT GCCAGCCGGC AAGACCCCGG
         CAGACAATAT AGAAACTGCC GACCTTTCGG CAAGCGTTCC CACCCGCCCT
    151 GCCGAACCCG AAAGAAAAAC GCTGGCAGAT TACGGCGGCT ACCCGTCCGC
    201 ACTGGATGCA GTGAAACAGA AAAACGATGC CGCCGTCGCC GCCTATTTGG
    251 AAAACGCCGG CGACAGCGCG ATGGCGGAAA ATGTCCGCAA CGAGTGGCTG
    301 AAGTCTTTGG GCGCACGCAG ACAGTGGACG CTGTTTGCAC AGGAATACGC
    351
         CAAACTCGAA CCGGCAGGGC GCGCCCAAGA AGTCGAATGC TACGCCGATT
         CGAGCCGCAA CGACTATACG CGTGCCGCTG AACTGGTCAA AAATACGGGC
        AAACTGCCTT CGGGCTGCAC CAAACTGTTG GAACAGGCAG CCGCATCCGG
    501
        CTTGTTGGAC GGCAACGACG CCTGGAGGCG CGTGCGCGGA CTGCTGGCCG
    551 GCCGCCAAAC CACAGACGCA CGCAACCTTG CCGCCGCATT GGGCAGCCCG
    601 TTTGACGGCG GTACACAAGG TTCGCGCGAA TATGCCCTGT TGAACGTCAT
    651 CGGCAAAGAA GCACGCAAAT CGCCGAATGC CGCCGCCCTG CTGTCCGAAA
    701 TGGAAAGCGG TTTAAGCCTC GAACAACGCA GTTTCGCGTG GGGCGTATTG
    751 GGGCATTATC AGTCGCAAAA CCTCAATGTG CCTGCCGCCT TGGACTATTA
    801 CGGCAAGGTT GCCGACCGCC GCCAACTGAC CGACGACCAA ATCGAGTGGT
    851 ACGCCCGCG CGCCTTGCGC GCCCGACGTT GGGACGAGCT GGCCTCCGTT
    901 ATCTCGCATA TGCCCGAAAA ACTGCAAAAA AGCCCGACCT GGCTCTACTG
    951 GCTGGCACGC AGCCGCGCCG CAACGGGCAA CACGCAAGAG GCGGAAAAAC
   1001 TTTACAAACA GGCGGCAGCG ACGGGCAGGA ATTTTTATGC GGTGCTGGCA
   1051 GGGGAAGAAT TGGGTCGGAA AATCGATACG CGCAACAATG TGCCCGATGC
         CGGCAAAAAC AGCGTCCGCC GCATGGCGGA AGACGGTGCA GTCAAACGCG
        CACTGGTACT GTTCCAAAAC AGCCAATCTG CCGGTGATGC AAAAATGCGC
   1201 CGTCAGGCTC AGGCGGAATG GCGTTTTGCC ACACGCGGCT TTGACGAAGA
   1251 CAAGCTGCTG ACCGCCGCG AAACCGCGTT CGACCACGGT TTTTACGATA
   1301 TGGCGGTCAA CAGCGCGGAA CGCACCGACC GCAAACTCAA CTACACCTTG
   1351 CGCTATATTT CGCCGTTTAA AGACACGGTA ATCCGCCACG CGCAAAATGT
   1401 TAATGTCGAT CCGGCTTGGG TTTATGGGCT GATTCGTCAG GAAAGCCGCT
   1451 TCGTTATAGG CGCGCAATCC CGCGTAGGCG CGCAGGGGCT GATGCAGGTT
   1501 ATGCCTGCCA CCGCGCGCA AATCGCCGC AAAATCGGTA TGGATGCCGC
   1551 ACAACTTTAC ACCGCCGACG GG...
```

This corresponds to the amino acid sequence <SEQ ID 108; ORF 019>:

m019.pep (partial)

1 MYLPSMKHSL PLLAALVLAA CSSTNTLPAG KTPADNIETA DLSASVPTRP

BNSDOCID: <WO___9957280A2_I_>

```
51 AEPERKTLAD YGGYPSALDA VKQKNDAAVA AYLENAGDSA MAENVRNEWL
101 KSLGARRQWT LFAQEYAKLE PAGRAQEVEC YADSSRNDYT RAAELVKNTG
151 KLPSGCTKLL EQAAASGLLD GNDAWRRVRG LLAGRQTTDA RNLAAALGSP
201 FDGGTQGSRE YALLNVIGKE ARKSPNAAAL LSEMESGLSL EQRSFAWGVL
251 GHYQSQNLNV PAALDYYGKV ADRRQLTDDQ IEWYARAALR ARRWDELASV
301 ISHMPEKLQK SPTWLYWLAR SRAATGNTQE AEKLYKQAAA TGRNFYAVLA
351 GEELGRKIDT RNNVPDAGKN SVRRMAEDGA VKRALVLFQN SQSAGDAKMR
401 RQAQAEWRFA TRGFDEDKLL TAAQTAFDHG FYDMAVNSAE RTDRKLNYTL
451 RYISPFKDTV IRHAQNVNVD PAWVYGLIRQ ESRFVIGAQS RVGAQGLMQV
501 MPATAREIAG KIGMDAAQLY TADG...
```

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 109>:

```
ATGTACCCAC CCTCTCTGAA GCATTCCCTG CCGCTGCTGG TGGNCCTGGT
   1
      GCTTGCCGCG TGTTCTTNGA CAAACACACT GTCAGCCGAC AAGACCCCGG
   51
 101 CAGACAATAT AGAAACTGCC GACCTTTCGG CAAGCGTTCC CACCNGCCCT
 151 GCCGAACCCG AANGAAAAAC GTNGGCAGAT TACGGCGGCT ACCCGTCCGC
  201 ACTGGATGCA GTGAAACAGA AAAACGATGC CGCCGTCGCC GCCTATTTGG
 251 AAAACGCCGG CGACAGCGCG ATGGCGGAAA ATGTCCGCAA CGAGTGGCTG
      AAGTCTTTGG GCGCGCGCAG ACAGTGGACG CTGTNTGCAC ANGAATATGC
      NAAACTCGAA CCGGCANGGC GCGCCCAAGA AGTCGAATGC TACGCCGATT
  351
      CGAGCCGCAA CGACTATACG CGTGCCGCCG AACTGGTCAA AAATACGGGC
 451 AAACTGCCTT CGGGCTGCAC CAAACTGTTG GAACAGGCAG CCGCATCCGG
 501 CTTGTTGGAC GGCAACGACG CCTGGAGGCG CGTGCGCGGA CTGCTGGCCG
      GCCGCCAAAC CACAGACGCA CGCAACCTTG CCGCCGCATT GGGCAGCCCG
 551
 601 TTTGACGGCG GTACACAAGG TTCGCGCGAA TATGCCCTGT TGAACGTCAT
     CGGCAAAGAA GCACGCAAAT CGCCGAATGC CGCCGCCCTG CTGTCCGAAA
 651
 701 TGGAAAGCGG TTTAAGCCTC GAACAACGCA GTTTCGCGTG GGGCGTATTG
     GGGCATTATC AGTCGCAAAA CCTCAATGTG CCTGCCGCCT TGGACTATTA
 751
      NGGCAAGGTT GCCGACCGCC GCCAACTGAC CGACGACCAA ATCGAGTGGT
 801
 851
      ACGCCCGCGC CGCNNTNNGC NNNCGNNGTT NGNANGANNT GGCNNCCGNN
 901 ANCNCGNNNN TGCNNGANAA ACNNNNNNAN AGNCNNANNT NGNTNNANTG
      NNTGGCACGC AGCCGCGCCG CNACGGGCAA CACGCAANAN GCGGANAAAC
 951
1001
      TNTACAAACA GGCGGCAGCA NCGGGCANGA ATTTTTATGC NGTGCTGNCN
1051
      GGGGAAGAGT TGGGGCGCAN AATCGATACG CGCAACAATG TGCCCGATGC
1101
      CGGCAAAANC AGCGTCCTCC GTATGGCGGA AGACGGCGCG ATTAAGCGCG
      CGCTGGTGCT GTTCCGAAAC AGCCGAACCG CCGGCGATGC GAAAATGCGC
1151
      CGTCNGGCTC AGGCGGAATG GCGTTTCGCC ACACGCGGCT TCGATGAAGA
1201
1251
      CAAGCTGCTG ACCGCCGCG AAACCGCGTT CGACCACGGT TTTTACGATA
      TGGCGGTCAA CAGCGCGGAA CGCACCGACC GCAAACTCAA CTACACCTTG
1301
      CGCTACATTT CGNNNNTNA NGACACGGTA ATCCGCCACG CGCAAAATGT
1351
     TAATGTCGAT CCGGCGTGGG TTTACGGGCT GATTCGTCAG GAAAGCCGCT
1401
     TCGTTATGGG CGCGCAATCC CGCGTAGGCG CGCAGGGGCT GATGCAGGTT
1451
     ATGCCTGCCA CCGCGCGCA AATCGCCGGC AAAATCGGTA TGGATGCCGC
1501
     ACAACTTTAC ACCGCCGACG GCAATATCCG TATGGGGACG TGGTATATGG
1551
     CGGACACCAA ACGCCGCCTG CAAAACAACG AAGTCCTCGC CACCGCAGGC
1601
1651
     TATAACGCCG GTCCCGGCAG GGCGCGCCGA TGGCAGGCGG ACACGCCCCT
     CGAAGGCGCG GTATATGCCG AAACCATCCC GTTTTCCGAA ACGCGCGACT
1701
1751
     ATGTCAAAAA AGTGATGGCC AATGCCGCCT ACTACGCCTC CCTCTTCGGC
     GCGCCGCACA TCCCGCTCAA ACAGCGTATG GGCATTGTCC CCGCCCGCTG
1851
```

## This corresponds to the amino acid sequence <SEQ ID 110; ORF 019.a>:

· beb					
1	MYPPSLKHSL	PLLVXLVLAA	CSXTNTLSAD	KTPADNIETA	DLSASVPTXP
51	AEPEXKTXAD	YGGYPSALDA	VKOKNDAAVA	AYLENAGDSA	MARNIJONEUT
101	KSTCAKKOMI	LXAXEYAKLE	PAXRAOEVEC	YADSSRNDVT	DAACTIMO
151	KLPSGCTKLL	EQAAASGLLD	GNDAWRRVRG	I.I.ACROTTDA	DATTANTOOD
201	FDGGTQGSRE	YALLNVIGKE	ARKSPNAAAI.	LSEMESCIST	PODCENTACTO
251	GHIQSQNLNV	PAALDYXGKV	ADRROLTDDO	TEWYADAAVV	VDVVVVVVV
301	AAAAAXKXXX	XXXXXXXXX	SRAATGNTOX	AYKI YKONNN	VOWNERSTOR
351	GEFFGKYIDI	RNNVPDAGKX	SVLRMAEDGA	TKRALVI FDM	CDMACDAMAD
401	RXAQAEWRFA	TRGFDEDKLL	TAAQTAFDHG	FYDMAVNSAE	RTDRKINVTI
					WI DIVIVINITIE

451	RYISXXXDTV	IRHAQNVNVD	PAWVYGLIRQ	ESRFVMGAQS	RVGAQGLMQV
501	MPATABETAG	KTCMDAAOT.V	TARCNIPMOT	MAMAULAMAM	ONNEULADAC

501 MPATAREIAG KIGMDAAQLY TADGNIRMGT WYMADTKRRL QNNEVLATAG 551 YNAGPGRARR WQADTPLEGA VYAETIPFSE TRDYVKKVMA NAAYYASLFG 601 APHIPLKQRM GIVPAR*

### m019/a019 88.9% identity over a 524 aa overlap

m019.pep	10 MYLPSMKHSLPLLAA      :      : MYPPSLKHSLPLLVX 10	1111111	111   1111	1111111111	1111 1111	1 11 11
m019.pep	70 YGGYPSALDAVKQKN             YGGYPSALDAVKQKN 70	111111111	111111111	1111111111	THEFT I	
m019.pep	130 PAGRAQEVECYADSS			111111111	1111111111	
m019.pep	190 LLAGROTTDARNLAA            LLAGROTTDARNLAA 190	11111111	111111111	111111111	111111111	
m019.pep	250 EQRSFAWGVLGHYQS               EQRSFAWGVLGHYQS 250		11 11 11 11 11		1111 1	
m019.pep	310 ISHMPEKLQKSPTWL I: XXXXXXKXXXXXXXXXXXXXXXXXXXXXXXXXXXXX	11111	11111   111	111111:1	11111 1111	
m019.pep	370 RNNVPDAGKNSVRRM           RNNVPDAGKXSVLRM 370	1111:111	1111:11:11	1111111111	1111111111	111111
m019.pep	430 TAAQTAFDHGFYDMA            TAAQTAFDHGFYDMA 430		111111111	1111111	İ HILLI	111111
m019.pep	490 ESRFVIGAQSRVGAQ      :        ESRFVMGAQSRVGAQ 490		11111111111		NIRMGTWYMA 530	ADTKRRL 540
a019	QNNEVLATAGYNAGPO 550	GRARRWQAD 560	FPLEGAVYAET 570	TIPFSETRDY 580	VKKVMANAAY 590	YASLFG 600

Computer analysis of this amino acid sequence gave the following results: Homology with a predicted ORF from N. gonorrhoeae

ORF 019 shows 95.5% identity over a 89 aa overlap with a predicted ORF (ORF 019.ng) from N. gonorrhoeae:

g019/m019

```
20
                                                 40
g019.pep
                   LLAALVLAACSSTNTLPAGKTPADNIETADLSASVPTRPAEPEGKTLAD
                   MYLPSMKHSLPLLAALVLAACSSTNTLPAGKTPADNIETADLSASVPTRPAEPERKTLAD
m019
                        20
                                30
                                        40
                                                50
         50
                 60
                         70
                                80
g019.pep
          YGGYPSALDAVKQNNDAAAAAYLENAGDSAMAENVRKEWL
          m019
          YGGYPSALDAVKQKNDAAVAAYLENAGDSAMAENVRNEWLKSLGARRQWTLFAQEYAKLE
                        80
                                90
                                       100
```

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 111>: g023.seq

- ATGGTAGAAC GTAAATTGAC CGGTGCCCAT TACGGTTTGC GCGATTGGGT 1
  - 51 AATGCAGCGT GCGACTGCGG TTATTATGTT GATTTATACC GTTGCACTTT
  - 101 TAGTGGTTCT ATTTGCCCTG CCTAAAGAAT ATCCGGCATG GCAGGCATTT
  - TTTAGTCAAG CTTGGGTAAA AGTATTTACC CAAGTGAGCT TTATCGCCGT
  - 201 ATTCTTGCAC GCTTGGGTGG GTATCCGCGA TTTGTGGATG GACTATATCA
  - 251 AACCCTTCGG CGTGCGTTTG TTTTTGCAGG TTGCCACCAT TGtctGGCTG
  - 301 GTCGGCTGCC TCGTGTATTC AGTTAAAGTG ATTTGGGGGT AA

### This corresponds to the amino acid sequence <SEQ ID 112; ORF 023.ng>: g023.pep

- MVERKLTGAH YGLRDWVMQR ATAVIMLIYT VALLVVLFAL PKEYPAWQAF 1
- FSQAWVKVFT QVSFIAVFLH AWVGIRDLWM DYIKPFGVRL FLQVATIVWL 51
- 101 VGCLVYSVKV IWG*

## The following partial DNA sequence was identified in N. meningitidis <SEQ ID 113>:

m023.seq

- ATGGTAGAAC GTAAATTGAC CGGTGCCCAT TACGGTTTGC GCGATTGGGT 1 GATGCAACGT GCGACTGCGG TTATTATGTT GATTTATACC GTTGCACTTT 51 101 TAGTGGTTCT ATTTTCCCTG CCTAAAGAAT ATTCGGCATG GCAGGCATTT 151 TTTAGTCAAA CTTGGGTAAA AGTATTTACC CAAGTGAGCT TCATCGCCGT 201 ATTCTTGCAC GCTTGGGTGG GTATCCGCGA TTTGTGGATG GACTATATCA 251 AACCCTTCGG CGTGCGTTTG TTTTTGCAGG TTGCCACCAT CGTTTGGCTG
- 301 GTCGGCTGTC TCGTGTATTC AGTTAAAGTG ATTTGGGGGT AA

### This corresponds to the amino acid sequence <SEQ ID 114; ORF 023>: m023.pep

- MVERKLTGAH YGLRDWVMQR ATAVIMLIYT VALLVVLFSL PKEYSAWQAF 1
  - FSQTWVKVFT QVSFIAVFLH AWVGIRDLWM DYIKPFGVRL FLQVATIVWL 101 VGCLVYSVKV IWG*

### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 115>: a023.seq

- ATGGTAGAAC GTAAATTGAC CGGTGCCCAT TACGGTTTGC GGGATTGGGC 1
  - 51 GATGCAACGT GCGACCGCGG TTATTATGTT GATTTATACC GTTGCACTTT 101
  - TAGTGGTTCT ATTTGCTCTG CCTAAAGAAT ATTCGGCATG GCAGGCATTT TTTAGTCAAA CTTGGGTAAA AGTATTTACC CAAGTGAGCT TCATCGCCGT

201	ATTCTTGCAC	GCTTGGGTGG	GTATCCGCGA	TTTGTGGATG	GACTATATNA
251	AACCCTTCGG	CGTGCGTTTG	TTTTTGCAGG	TTGCCACCAT	CGTCTGGCTG

301 GTCGGCTGCT TGGTGTATTC AATTAAAGTA ATTTGGGGGT AA

This corresponds to the amino acid sequence <SEQ ID 116; ORF 023.a>:

a023.pep

- 1 MVERKLTGAH YGLRDWAMQR ATAVIMLIYT VALLVVLFAL PKEYSAWQAF
- 51 FSQTWVKVFT QVSFIAVFLH AWVGIRDLWM DYXKPFGVRL FLQVATIVWL
- 101 VGCLVYSIKV IWG*

m023/a023 96.5% identity over a 113 aa overlap

```
40
                                                  60
          MVERKLTGAHYGLRDWVMQRATAVIMLIYTVALLVVLFSLPKEYSAWQAFFSQTWVKVFT
m023.pep
          MVERKLTGAHYGLRDWAMQRATAVIMLIYTVALLVVLFALPKEYSAWQAFFSQTWVKVFT
a023
                10
                       20
                               30
                                      40
                70
                       80
                               90
                                      100
                                             110
          QVSFIAVFLHAWVGIRDLWMDYIKPFGVRLFLQVATIVWLVGCLVYSVKVIWGX
m023.pep
          a023
          QVSFIAVFLHAWVGIRDLWMDYXKPFGVRLFLQVATIVWLVGCLVYSIKVIWGX
                70
                       80
                               90
                                      100
                                             110
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae

ORF 023 shows 97.3% identity over a 113 aa overlap with a predicted ORF (ORF 023.ng) from N. gonorrhoeae:

g023/m023

```
20
                               3.0
                                       40
                                              50
                                                      60
g023.pep
          MVERKLTGAHYGLRDWVMQRATAVIMLIYTVALLVVLFALPKEYPAWQAFFSQAWVKVFT
          MVERKLTGAHYGLRDWVMQRATAVIMLIYTVALLVVLFSLPKEYSAWQAFFSQTWVKVFT
m023
                10
                       20
                               30
                                       40
                                              50
                                                      60
                               90
                                      100
g023.pep
          QVSFIAVFLHAWVGIRDLWMDYIKPFGVRLFLQVATIVWLVGCLVYSVKVIWGX
          m023
          QVSFIAVFLHAWVGIRDLWMDYIKPFGVRLFLQVATIVWLVGCLVYSVKVIWGX
                70
                       80
                               90
                                      100
```

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 117>: g025.seq

```
ATGTTGAAAC AAAcgACACT TTTGGCAGCT TGTACCGCCG TTGCCGCTCT
  1
     GTTGGGCGGT TGcgCCACCC AACAGCCTGC TccTGTCATT GCAGGCAATT
 51
     CAGGTATGCA GACCGTATCG TCTGCGCCGG TTTACAATCC TTATGGCGCA
101
     ACGCCGTACA ATGCCGCTCC TGCCGCCAac gatgcGCCgT ATGTGCCGCC
     CGTGCAAact gcgccggttT ATTCGCCTCC TGCTTATGTT CCGCcgtCTG
251
     CACCTGCCGT TTCGGgtaca tatgtTCCTT CTTACGCACC CgtcgACATC
     aacgCGGCGa cgCataCTAT TGTGCGTGGC GACACGGtgt acaACATTTc
301
     caaAcgCtac CATATCTCTC AAGACGATTT CCGTGCGTGG AACGGCATGA
     CCGACAATAC GTTGAGCATC GGTCAGATTG TTAAAGTCAA ACCGGCaggA
401
     TATGCCGCAC CGAAAACCGC AGCCGTAGAA AGCAGGCCCG CCGTACCGGC
451
     TGCCGCGCAA ACCCCTGTGA AACCCGCCGC gcaACCGCCC GTTCAGTCCG
551
     CGCCGCAACC TGCCGCGCCC GCTGCGGAAA ATAAAGCGGT TCCCGCCCCC
601 GCGCCCGCCC CGCAATCTCC TGCCGCTTCG CCTTCCGGCA CGCGTTCGGT
    CGGCGGCATT GTTTGGCAGC GTCCGACCCA AGGTAAAGTG GTTGCCGATT
```

BNSDOCID: <WO __ 9857280A2 | >

```
701 TCGGCGGCG CAACAAGGGT GTCGATATTG CCGGCAATGC CGGACAACCC
751 GTTTTGGCGG CGGCTGACGG CAAAGTGGTT TATGCCGGTT CAGGTTTGAG
801 GGGATACGGA AACTTGGTCA TCATCCAGCA CAATTCCTCT TTCCTGACCG
851 CGTACGGGCA CAACCAAAAA TTGCTGGTCG GCGAAGGTCA GCAGGTCAAA
901 CGCGGTCAGC AGGTTGCTTT GATGGGTAAT ACCGATGCTT CCAGAACGCA
951 GCTTCATTTC GAGGTGCGTC AAAACGGCAA ACCGGTTAAC CCGAACAGCT
```

## This corresponds to the amino acid sequence <SEQ ID 118; ORF 025.ng>: g025.pep

```
1 MLKQTTLLAA CTAVAALLGG CATQQPAPVI AGNSGMQTVS SAPVYNPYGA
51 TPYNAAPAAN DAPYVPPVQT APVYSPPAYV PPSAPAVSGT YVPSYAPVDI
101 NAATHTIVRG DTVYNISKRY HISQDDFRAW NGMTDNTLSI GQIVKVKPAG
151 YAAPKTAAVE SRPAVPAAAQ TPVKPAAQPP VQSAPQPAAP AAENKAVPAP
201 APAPQSPAAS PSGTRSVGGI VWQRPTQGKV VADFGGGNKG VDIAGNAGQP
251 VLAAADGKVV YAGSGLRGYG NLVIIQHNSS FLTAYGHNQK LLVGEGQQVK
301 RGQQVALMGN TDASRTQLHF EVRQNGKPVN PNSYIAF*
```

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 119>:

```
..GTGCCGCCGG TGCAAAGCGC GCCGGTTTAT ACGCCTCCTG CTTATGTTCC
   1
        GCCGTCTGCA CCTGCCGTTT CGGGTACATA CGTTCCTTCT TACGCACCCG
  51
        TCGACATCAA CGCGGCGACG CATACTATTG TGCGCGGCGA CACGGTGTAC
 101
        AACATTTCCA AACGCTACCA TATCTCTCAA GACGATTTCC GTGCGTGGAA
 151
        CGGCATGACC GACAATACGT TGAGCATCGG TCAGATTGTT AAAGTCAAAC
 201
        CGGCAGGATA TGCCGCACCG AAAGCCGCAG CCGTAAAAAG CAGGCCCGCC
 251
        GTACCGGCTG CCGCGCAACC GCCCGTACAG TCCGCACCCG TCGACATTAA
 301
        CGCGGCGACG CATACTATTG TGCGCGGCGA CACGGTGTAC AACATTTCCA
 351
        AACGCTACCA TATCTCTCAA GACGATTTCC GTGCGTGGAA CGGCATGACC
 401
        GACAATATGT TGAGCATCGG TCAGATTGTT AAAGTCAAAC CGGCAGGATA
 451
 501
        TGCCGCACCG AAAACCGCAG CCGTAGAAAG CAGGCCCGCC GTACCGGCTG
        CCGTGCAAAC CCCTGTGAAA CCCGCCGCG AACCGCCTGT GCAGTCCGCG
 551
        CCGCAACCTG CCGCGCCCGC TGCGGAAAAT AAAGCGGTTC CCGCGCCCGC
 651
        CCCGCAATCT CCTGCCGCTT CGCCTTCCGG CACGCGTTCG GTCGGCGGCA
        TTGTTTGGCA GCGTCCGACG CAAGGTAAAG TGGTTGCCGA TTTCGGCGGC
 701
        AACAACAAGG GTGTCGATAT TGCCGGTAAT GCGGGACAGC CCGTTTTGGC
 751
        GGCGGCTGAC GGCAAAGTGG TTTATGCCGG TTCAGGTTTG AGGGGATACG
 801
        GAAACTTGGT CATCATCCAG CATAATTCTT CTTTCCTGAC CGCATACGGG
 851
        CACAACCAAA AATTGCTGGT CGGCGAGGGG CAGCAGGTCA AACGCGGTCA
 901
        GCAGGTTGCT TTGATGGGCA ATACCGATGC TTCCAGAACG CAGCTTCATT
 951
1001
        TCGAGGTGCG TCAAAACGGC AAACCGGTTA ACCCGAACAG CTATATCGCG
1051
        TTCTGA
```

### This corresponds to the amino acid sequence <SEQ ID 110; ORF 025>:

```
m025.pep (partial)

1 ..VPPVQSAPVY TPPAYVPPSA PAVSGTYVPS YAPVDINAAT HTIVRGDTVY
51 NISKRYHISQ DDFRAWNGMT DNTLSIGQIV KVKPAGYAAP KAAAVKSRPA
101 VPAAAQPPVQ SAPVDINAAT HTIVRGDTVY NISKRYHISQ DDFRAWNGMT
151 DNMLSIGQIV KVKPAGYAAP KTAAVESRPA VPAAVQTPVK PAAQPPVQSA
201 PQPAAPAAEN KAVPAPAPQS PAASPSGTRS VGGIVWQRPT QGKVVADFGG
251 NNKGVDIAGN AGQPVLAAAD GKVVYAGSGL RGYGNLVIIQ HNSSFLTAYG
301 HNQKLLVGEG QQVKRGQQVA LMGNTDASRT QLHFEVRQNG KPVNPNSYIA
```

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 111>:

1 ATGTTGACAC CAACAACACT TTAGGTAGCT TGTACCGCCC TTGCCGCTCA
51 GTTGGGCGGA TGCCCCACCC AACACCCTTC TCCTGTCATT GCAGGCAATT
101 CAGGTATGCA GACCGTACCG TCTGCGCCGG TTTACAATCC TTATGGCGCA

1.51	
151	ACGCCGTACA ATGCCGCTCC TGCCGCCAAC GATGCGCCGT ATGTGCCGCC
201	GGTGCAAAGC GCGCCGGTTT ATANGCCTCC TGCTTATGTT CCGCCGTCTG
251	CACCTGCCGT TTCGGGTACA TACGTTCCTT CTTACGCANC CGTCGACATC
301	AACGCGGCGA CCCATACTAT TGTGCGCGGC GACACCGTGT ACAAGATTTC
351	CAAATGCTAC CATATCTCTC AAGACGATTT CCGTGCGTGG AACGGCATGA
401	CCGACAATAC GTTGAGCATC GGTCAGATTG TTAAAGTCAA ACCGGCAGGA
451	TATGCCGCAC CGAAAGCCGC AGCCGTAAAA AGCAGGCCCG CCGTACCGGC
501	TGCCGCGCAA CCGCTCGTAC AGTCCGCACC CGTCGACATC AACGCGGCGA
551	CGCATACTAT TGTGCGCGGC GACACGGTGT ACAACATTTC CAAACGCTAC
601	CATATCTCTC AAGACGATTT CCGTGCGTGG AACGGCATGA CCGACAATAC
651	GTTGAGCATC GGTCAGATTG TTAAAGTCAA ACCGGCAGGA TATGCCGCAC
701	CGAAAGCCGC AGCCGTAAAA AGCAGGCCCG CCGTACCGGC TGCCGTGCAA
751	ACCCCTGTGA AACCCGCCGC GCAACCGCCT GTGCAGTCCG CGCCGCAACC
	TGCCGCGCCC GCTGCGGAAA ATAAAGCGGT TCCCGCGCCC GCCCCGCAAT
801	
851	CTCCTGCCGC TTCGCCTTCC GGCACGCGTT CGGTCGGCGG CATTGTTTGG
901	CAGCGTCCGA CGCAAGGTAA AGTGGTTGCC GATTTCGGCG GCAACAACAA
951	GGGTGTCGAT ATTGCAGGAA ATGCGGGACA GCCCGTTTTG GCGGCGGCTG
1001	ACGGCAAAGT GGTTTATGCA GGTTCCGGTT TGAGGGGATA CGGCAATTTG
1051	GTCATCATCC AGCATAATTC TTCCTTCCTG ACCGCATACG GGCACAACCA
1101	AAAATTGCTG GTCGGCGAAG GCCAGCAGGT CAAACGCGGG CAGCAGGTCG
1151	CTTTGATGGG CAATACCGAG GCTTCTAGAA CGCAGCTTCA TTTCGAGGTG
1201	CGGCAAAACG GCAAACCGGT TAATCCGAAC AGCTATATCG CGTTCTGA
This correspond	s to the amino acid sequence <seq 025.a="" 112;="" id="" orf="">:</seq>
a025.pep	2 to the minute solution of 22 to 112, 314 325.0°.
	MITTOTH + 1/2 CT21 2 A OLC COTTOURS DUT 2 CHECKOTHE CA DUNNING
1	MLTPTTL*VA CTALAAQLGG CPTQHPSPVI AGNSGMQTVP SAPVYNPYGA
51	TPYNAAPAAN DAPYVPPVQS APVYXPPAYV PPSAPAVSGT YVPSYAXVDI
101	NAATHTIVRG DTVYKISKCY HISQDDFRAW NGMTDNTLSI GQIVKVKPAG
151	YAAPKAAAVK SRPAVPAAAQ PLVQSAPVDI NAATHTIVRG DTVYNISKRY
201	HISQDDFRAW NGMTDNTLSI GQIVKVKPAG YAAPKAAAVK SRPAVPAAVQ
251	TPVKPAAQPP VQSAPQPAAP AAENKAVPAP APQSPAASPS GTRSVGGIVW
301	QRPTQGKVVA DFGGNNKGVD IAGNAGQPVL AAADGKVVYA GSGLRGYGNL
351	VIIQHNSSFL TAYGHNQKLL VGEGQQVKRG QQVALMGNTE ASRTQLHFEV
401	RQNGKPVNPN SYIAF*
	97.4% identity over a 351 aa overlap
m025/a025	
m025/a025	
m025/a025	10 20 30
	10 20 30
m025/a025	10 20 30 VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS
m025.pep	10 20 30  VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS  !!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
	10 20 30  VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS          :
m025.pep	10 20 30 VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS !!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
m025.pep	10 20 30  VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS           :
m025.pep a025	10 20 30  VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS
m025.pep	10 20 30  VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS           :
m025.pep a025 m025.pep	10 20 30  VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS
m025.pep a025	10 20 30  VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS
m025.pep a025 m025.pep	10 20 30  VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS
m025.pep a025 m025.pep	10 20 30  VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS
m025.pep a025 m025.pep a025	10 20 30  VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS
m025.pep a025 m025.pep	10 20 30  VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS
m025.pep a025 m025.pep a025	10 20 30  VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS
m025.pep a025 m025.pep a025 m025.pep a025	10 20 30  VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS
m025.pep a025 m025.pep a025 m025.pep a025	10 20 30  VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS
m025.pep a025 m025.pep a025 m025.pep a025	10 20 30  VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS
m025.pep a025 m025.pep a025 m025.pep a025	10
m025.pep a025 m025.pep a025 m025.pep a025	10
m025.pep a025 m025.pep a025 m025.pep a025	10

a025	KAVPAPAPQSP 280	AASPSGTRSVO	GIVWQRPTQ	GKVVADFGGNI 310		
		200	300	310	320	330
	280	290	300	310	320	330
m025.pep	GKVVYAGSGLRO	GYGNLVIIQHN	SSFLTAYGH	NQKLLVGEGQQ	QVKRGQQVALI	MGNTDASRT
a025	GKVVYAGSGLRO	!				1111.1111
	340	350	360	370	380	390
	340	350				
m025.pep	QLHFEVRQNGKE	PVNPNSYIAFX				
2025						
a025	QLHFEVRQNGKF 400	VNPNSYIAFX 410				

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae

ORF 025 shows 75.6% identity over a 353 aa overlap with a predicted ORF (ORF 025.ng) from N. gonorrhoeae:

m025/g025

m025.pep VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS GMQTVSSAPVYNPYGATPYNAAPAANDAPYVPPVQTAPVYSPPAYVPPSAPAVSGTYVPS g025 YAPVDINAATHTIVRGDTVYNISKRYHISQDDFRAWNGMTDNTLSIGQIVKVKPAGYAAP m025.pep g025 YAPVDINAATHTIVRGDTVYNISKRYHISQDDFRAWNGMTDNTLSIGQIVKVKPAGYAAP KAAAVKSRPAVPAAAQPPVQSAPVDINAATHTIVRGDTVYNISKRYHISQDDFRAWNGMT m025.pep g025  ${\tt DNMLSIGQIVKVKPAGYAAPKTAAVESRPAVPAAVQTPVKPAAQPPVQSAPQPAAPAAEN}$ m025.pep g025 --TAAVESRPAVPAAAQTPVKPAAQPPVQSAPQPAAPAAEN  ${\tt KAVPAPAP--QSPAASPSGTRSVGGIVWQRPTQGKVVADFGGNNKGVDIAGNAGQPVLAA}$ m025.pep KAVPAPAPAPQSPAASPSGTRSVGGIVWQRPTQGKVVADFGGGNKGVDIAGNAGQPVLAA g025  $\verb"ADGKVVYAGSGLRGYGNLVIIQHNSSFLTAYGHNQKLLVGEGQQVKRGQQVALMGNTDAS"$ m025.pep g025 ADGKVVYAGSGLRGYGNLVIIQHNSSFLTAYGHNQKLLVGEGQQVKRGQQVALMGNTDAS m025.pep RTQLHFEVRQNGKPVNPNSYIAFX

PCT/US99/09346



The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 113>: g031.seq

```
ATGGTGTCCC TCCGCTTCAG ATTCGGCAAC CACTTTAAAC GCCGACATTC
  1
 51
     TGACAATTTC CTTTTCCGCC AGCCAAATAT CATGCGTATC TTTCGGTTCG
101 GGCTTGTTGG GCATGGCAAC CTTCAACAGC CGCGCCATCA CAGGAATCGT
151
     CGTTCCCTGA ATCAGCAGCG ACAGCACCAC CACGGCAAAC GCCACATCAA
201
     ACAGCAGGTG CGAATTGGGA ACGCCCATCA CCAGCGGCAT CATCGCCAGC
    GAAATCGGTA CGGCTCCTCG CAAGCCCAAC CAACTGATAT ACGCCTTTTC
251
301 ACGCAGGCTG TAATTGAATT TCCACAAACC GCCGAACACT GCCAGCGGAC
     GCGCGACCAG CATCAGGAAC GCCGCAATCG CCAAGGCTTC CGCCGCCCTG
401 TCCAACACGC CGGCGGAGA AACCAGCAGA CCGAGCATGA CGAACAAAGT
451
     TGCCTGCGCC AGCCAAGCCA AACCGTCCAT CACACGCAAA ACGTGTTCCG
     TCGCACGGTT GCGCTGGTTA CCGACAATGA TGCCGGCAAG GTAAACCGCC
551 AAAAAGCCGC TGCCGCCTAT GGTATTGGTA AACGCAAACA CAAGCAGCCC
601 GCCCGACACA ATCATCAGCG CGTACAGACC TTCCGtacac acctccaatt
651 cccaatcaac gtcatagctg tctcccgtgt taaaatgttc ttcacttcag
701 aatcccccc ttcttcccag cccgaaacct tcatgtgtta naccctgggg
751 tgccccaacg gatttagtaa cctcccaatg actergerry regected
801 cgcccgcttt ctccttccgg gaaaacttgt tgtccccgtc ttacattaa
    tgccccaacg gatttagtaa cctcccaatg actctgcttg tcgcccctt
```

## This corresponds to the amino acid sequence <SEQ ID 114; ORF 031.ng>: g031.pep

```
1 MVSLRFRFGN HFKRRHSDNF LFRQPNIMRI FRFGLVGHGN LQQPRHHRNR
51 RSLNQQRQHH HGKRHIKQQV RIGNAHHQRH HRQRNRYGSS QAQPTDIRLF
101 TQAVIEFPQT AEHCQRTRDQ HQERRNRQGF RRPVQHAGGR NQQTEHDEQS
151 CLRQPSQTVH HTQNVFRRTV ALVTDNDAGK VNRQKAAAAY GIGKRKHKQP
201 ARHNHQRVQT FRTHLQFPIN VIAVSRVKMF FTSESPPSSQ PETFMCXTLG
251 CPNGFSNLPM TLLVAPFARF LLPGKLVVPV LH*
```

## The following partial DNA sequence was identified in N. meningitidis <SEQ ID 115>: m031.seq (partial)

```
1 ...CGCCTGAAGC ACGGTGTCGG ACTGCATTTC TATTCGGCTA TACGCCTTTT
51 CACGCAGGCT GTAATTGAAT TTCCACAAAC CGCCGAACAC TGCCGACGGA
101 CGCGCGACCA GCATCAGGAA CGCCGCAATC GCCAAGGCTT CCGCCGCCCT
151 GTCCAACACG TTGGCAGGAG AAACCAGCAG CAAAGGCATT CCCAAACGTG
201 CGGACAAAGT GGTCGAAACC ACGCTCAGAA ACAACAGTGC GCCACCCGGC
251 AG....
```

## This corresponds to the amino acid sequence <SEQ ID 116; ORF 031>: m031.pep (partial)

```
1 ...RLKHGVGLHF YSAIRLFTQA VIEFPQTAEH CRRTRDQHQE RRNRQGFRRP
VQHVGRRNQQ QRHSQTCGQS GRNHAQKQQC ATRQ....
```

### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 117>:

```
a031.seq

1 ATACGCCTTT TCACGCAGGC TGTAATTGAA TTTCCACAAA CCGCCGAACA
51 CTGCCGGCGG ACGCCGCACC AGCATCAGGA ACGCCGCAAT CGCCAAGGCT
101 TCCGCCGCCC CGTCCAACAC GTTGGCAGGA GAAACCAGCA GCAAAGGCAT
151 TCCCAAACGT GCGGACAAAG TGGTCGAAAC CACGCTCAGA AACAACAGTG
201 CGCCACCCGG CAG
```

### This corresponds to the amino acid sequence <SEQ ID 118; ORF 031.a>:

```
a031.pep (partial)

1 IRLFTQAVIE FPQTAEHCRR TRDQHQERRN RQGFRRPVQH VGRRNQQQRH
51 SQTCGQSGRN HAQKQQCATR Q
```

BNSDOCID: <WO__9957280A2 J_>

### m031/a031 100.0% identity over a 71 aa overlap

m031.pep	10 RLKHGVGLHFYSAI	20	30 POTAFUCEDTED	40	50	60
				QAQERRNRQQ 	FRRPVQHVG	RRNQQ
a031	I	RLFTQAVIEF	POTAEHCRRTRD	QHQERRNROG	FRRPVOHVG	RRNOO
		10	20	30	40	
	70	80				
m031.pep	QRHSQTCGQSGRNH					
0.04		111111111				
a031	QRHSQTCGQSGRNH	IAQKQQCATRQ				
	50 60	70				

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae

ORF 031 shows 60.0% identity over a 85 aa overlap with a predicted ORF (ORF 031.ng) from N. gonorrhoeae:

m031/g031

m031.pep			RT	10 KHGVGLH <b>F</b> YS	20	30
			112			
g031	NOOROHHHGKRH	IKQQVRIGNA	HHQRHHRQRN	::  : : RYGSSQAQPT	  DIRLFTQA	
	60	70	80	90	100	110
	40	50	60	70	80	1
m031.pep	CRRTRDQHQERR	NRQGFRRPVQ	HVGRRNQQQR	HS-QTCGQSGI	RNHAOKOOC	'ATRO
	1:111111111		1:1       :	1: 1:1 ::		
g031	CORTROQHQERR	NRQGFRRPVQI	HAGGRNQQTE	HDEQSCLRQPS	SOTVHHTON	VERRTVALV
	120	130	140	150	160	170
g031	TDNDAGKVNRQK	AAAAYGIGKRI	KHKQPARHNH	QRVOTFRTHLO	DEPINUTAV	SBUKMEETC
	180	190	200	210	220	230

# The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 119>:

1	ATGCGGCGAA	ACGTGCCTGC	CGTCGCCGTA	TTGCGCCGCC	САССАТТССА
51	GGCGTTTTTG	GATTTGGCGT	TGGCTCAGGC	GCGTGCCGTT	CCTGCCGGTA
101	AACAGGGCTT	TGCCGTCCGA	TGCCGTCTGA	CGCAGCGGCA	CATACTTTTT
151	CAAGGCTTCC	ACGCTTTTGC	CGGTCAGCGG	AACCTGACGC	TECTTECCCC
201	CTTTGCCGGT	AACGTGTACC	CACGCTTCGT	CCAAATATAC	ATCATCTCCA
251	TTCAAGCCGT	GTATCTCGCT	CACGCGCAAA	CCGCTGCCGT	ACATCACCTC
301	GAACAGCGCG	TGGTCGCGCA	CCGCCAGCGG	GTCGCCGCCG	TCCACGGGCA
351	AATCCAACAT	CCGGTTCAGC	CATTCCTGCG	GCAGGGCTTT	GGGTACGCCC
401	TCGGGCTGCT	TCGGCGGTTT	GATGTCGGCG	GTCGGGTCGG	CGCGCATCAC
451	CCCGCGTTTG	ACCAGCCAGG	CGCAATACTG	CCGCCACGCC	GACACCTTCC
501	GCGCCAGCGT	CCGACCGTCC	AAACCGCGCT	GCGACAGCCG	CCGCAACGCC
551	GCCGTAAAAT	CGCGCCGCGA	CAAGTCCTGC	GGCACGCcac	Ctacamemme
601	AGACGGCATT	TGTGCCAACA	GTGCAAACAG	TTCTTCCAAA	TOGOGOGOT
651	ATGCCGCAAC	CGTGTGCTCC	GACTTGCCCT	CGCGCACGAT	GTTTTCCAAA
701	TAAGCGTCAA	AATacgccgC	AAACccqTCC	AAAACCATAA	CCGTCCCACA
751	CAAATATCAA	AAAACCAGTG	A		JULI CCCACA

# This corresponds to the amino acid sequence <SEQ ID 110; ORF 032.ng>: g032.pep

1 MRRNVPAVAV LRRPRFEAFL DLALAQARAV PAGKQGFAVR CRLTQRQIVF 51 QGFHAFAGQR NLTLLAPFAG NVYPRFVQIY IICIQAVYLA HAQTAAVHQL 101 EQRVVAHRQR VAAVHGQIQH PVQPFLRQGF GYALGLLRRF DVGGRVGAHQ

```
PAFDQPGAIL PPRRQLARQR PTVQTALRQP PQRRRKIAPR QVLRHAACIF
          201
              RRHLCQQCKQ FFQIAPVCRN RVLRLALAHD VFQISVKIRR KPVQNHNRPT
              QISKNQ*
The following partial DNA sequence was identified in N. meningitidis <SEO ID 111>:
     m032.seq (partial)
              ATGCGGCGAA ACGTGCmTGC mGTCGCCGTT kTGCGCCGCC CATTGCGCCA
            1
           51
              AACGTTTTTG GATTTGGCGT TGGCTCAGGC GCGTGCCGTT CCTGCCGGTA
          101
              AACAGGGCTT TGCCGTCCGA TGCCGTCTGA CGCAGCGGCA GATAGTTTTT
          151 CAGGGCTTCC ACGCTTTTGC CGACCAGCGG CACCTGCCGC TgTT.GCGCC
          201 CTTTGCCGAT AACGTGTACC CACGCYTCGT CCAAATAGAC ATCATCTGCA
          251 TTCAAGCCGT GTATCTCGCT CACGCGCAAA CCGCTGCCGT ACATCAGTTC
          301 GAACAGGGCG TGGTCGCGCA CCGCCAGCGG GTCGCCGCCG TCCACGGGCA
          351
              AATCCAGCAT CCGGTTCAGC CATTCCTGCG GCAGGGCTTT GGGTACGCGC
              401
          451 GCCGCGCTTT ACCAGCCAAA CGCAATACTG CCGCCAAGAC GAAAGCTTGC
          501 GAGCCAGCGT CCGTTCCCCC AAACCGCG...
This corresponds to the amino acid sequence <SEQ ID 112; ORF 032>:
     m032.pep
              (partial)
              MRRNVXAVAV XRRPLRQTFL DLALAQARAV PAGKQGFAVR CRLTQRQIVF
              QGFHAFADQR HLPLXAPFAD NVYPRXVQID IICIQAVYLA HAQTAAVHQF
           51
          101
              EQGVVAHRQR VAAVHGQIQH PVQPFLRQGF GYALGLLRRF DVGGRVGVHQ
          151 AALYQPNAIL PPRRKLASQR PFPQTA...
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 113>:
     a032.seg
              ATGCGGCGAA ACGTGCCTGC CGTCGCCGTT TTGCGCCGCC CATTGCGCCA
          51
              AACGTTTTTG GATTTGGCGT TGGCTCAGGC GCGTGCCGTT CCTGCCGGTA
             AACAGGGCTT TGCCGTCCGA TGCCGTCTGA CGCAGCGGCA GATAGTTTTT
         151 CAGGGCTTCC ACGCTTTTGC CGGTCAGCGG AACCTGCCGC TGCTTGCGTC
             CTTTGCCGGT AACGTGTACC CACGCCTCGT CCAAATATAC ATCATCTGCA
         251
              TTCAAGCCGT GTATCTCGCT CACGCGCAAA CCGCTGCCGT ACATCAGTTC
              GAACAGCGCG TGATCGCGCA CCGCCAGCGG GTCGCCGCCG TCCACGGGCA
         351 AATCCAGCAT CCGGTTCAGC CATTCCTGCG GCAGGGCTTT GGGTACGCGC
         ACCGCGTTTG ACCAGCCAGG CGCAATACTG CCGCCAAGAC GACAGCTTGC
              GCGCCAGCGT CCGCGCATTC AAACCGCGCT GCGACAGCCG CCGCAACGCC
         501
              GCCGTAAAAT CGCGCTGCGA CAAGCCCTGC GGCACGCCGC CTGCATCTTC
              AGACGGCATT TGTGCCAACA GCGCAAACAG TTCTTCCAAA TCGCGCCGGT
         601
         651 ATGCCGCCAC CGTGTGCTCC GACTTGCCCT CGCGCACGAT GTTTTCCAAA
              TAAGCGTCAA AATGCGCCGC AAACCCGTCC AAAACCATAA CCGCCCCACA
         751 CAAATATCAA AAAAACAGTG A
This corresponds to the amino acid sequence <SEQ ID 114; ORF 032.a>:
     a032.pep
              MRRNVPAVAV LRRPLRQTFL DLALAQARAV PAGKQGFAVR CRLTQRQIVF
           1
              QGFHAFAGQR NLPLLASFAG NVYPRLVQIY IICIQAVYLA HAQTAAVHQF
          51
              EQRVIAHRQR VAAVHGQIQH PVQPFLRQGF GYALGLLRRF DVGGRVGMQQ
         151
              TAFDQPGAIL PPRRQLARQR PRIQTALRQP PQRRRKIALR QALRHAACIF
         201
              RRHLCQQRKQ FFQIAPVCRH RVLRLALAHD VFQISVKMRR KPVQNHNRPT
              QISKKQ*
            88.1% identity over a 176 aa overlap
m032/a032
                 MRRNVXAVAVXRRPLRQTFLDLALAQARAVPAGKQGFAVRCRLTQRQIVFQGFHAFADQR
    m032.pep
                 a032
                 {\tt MRRNVPAVAVLRRPLRQTFLDLALAQARAVPAGKQGFAVRCRLTQRQIVFQGFHAFAGQR}
                        10
                                  20
                                           30
                                                    40
```

m032.pep	70 HLPLXAPFADNVYP	80 RXVQIDIICI	90 QAVYLAHAQT	100 AAVHQFEQGV	110 VVAHRQRVAA	120 VHGQIQH
a032	NLPLLASFAGNVYP	RLVQIYIICI 80		'AAVHQFEQRV		VHGQIQH
	, •	80	90	100	110	120
	130	140	150	160	170	
m032.pep	PVQPFLRQGFGYAL	GLLRRFDVGG	RVGVHQAALY	QPNAILPPRE	KIA SOD DEDO	OTA
a032	PVQPFLRQGFGYAL	1     1   1   1   1   1	111::1:1:	11:11111	. 1 1 1 1 1	LLI
	130	140	<b>1</b> 50	160	170	180
a032	PORRRETATROATR	JAACT EDDUY	G005***			
	PQRRRKIALRQALRI	TAMCIFKKHL	CQQRKQFFQI.	APVCRHRVLR	LALAHDVFQI	SVKMRR
	190	200	210	220	230	240

Computer analysis of this amino acid sequence gave the following results: Homology with a predicted ORF from N. gonorrhoeae

ORF 032 shows 86.4% identity over a 176 aa overlap with a predicted ORF (ORF 032.ng) from *N. gonorrhoeae:* 

m032/g032

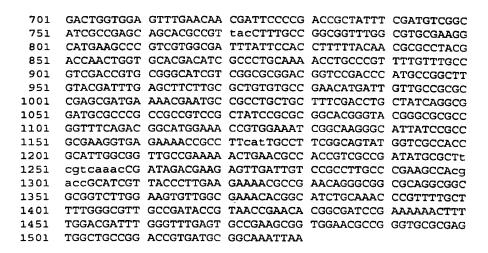
m032.pep	10 MRRNVXAVAVXRRP	::(	1   1   1   1   1   1   1	11111111	LILIBERTA	
g032	MRRNVPAVAVLRRP	RFEAFLDLA	LAQARAVPAGK	OGFAVRCRI	POROTVEOGE:	1111
	10	20	30	40	50	60
	70	80	90	100	110	100
m032.pep	HLPLXAPFADNVYP	RXVQIDIIC	CAVYLAHAOT	AAVHOFEOGY	WAMBOBINA	120
	• •   •   1   1   1   1   1   1   1   1	1				111111
g032	NLTLLAPFAGNVYPI	RFVQIYIICI	QAVYLAHAOT.	AAVHOLEORY	/VAHRORVAAT	MEOTOR
	70	80	90	100	110	120
	130	140	150	160	170	
m032.pep	PVQPFLRQGFGYALO	GLLRRFDVGG	RVGVHOAALY	OPNAILPPRE	T 7 O	יייי א
	_	1	111:11 1:	11.11111	-11 111 1	1.1
g032	PVQPFLRQGFGYAL	SLLRRFDVGG	RVGAHQPAFD	OPGAILPPRE	CLARORPTVC	1 1 2 T.D.O.D.
	130	140	150	160	170	180
g032	PQRRRKIAPRQVLRH	HAACI FRRHT	COOCKOFFOTZ	A DUCDMOUT D	T 7 T 7 !!D!!	
	190	200	210	220		
			210	220	230	240

# The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 115>:

_					
1	ATGGCGGCGG	CGGACAAACT	CTTGGGCGGC	GACCGCCGCA	GCGTCGCCAT
51	CATCGGAGAC	GGCGCGATGA	CGGCGGGGCA	GGCGTTTGAA	CCCTTCA A TO
101	GCGCGGGCGA	TATGGATGTG	GATTTGCTGG	TCGTCCTCAA	CGACAACGAA
151	ATGTCGATTT	CCCCCAACGT	CGGCGCGTTG	CCCDAATATC	TTGCCAGCAA
201	COTCOTOCO	CATATICCACO	Cacacacata	CCCAAAIAIC	TIGCCAGCAA
	2212213666	GATATGCACG	GACTGTTGAG	TACCGTCAAA	GCGCAAAcgg
251	GCAAGGTATT	AGACAAAATA	CCCGGCGCGA	TGGagtTTGC	CCANANACTC
301	GAACAcaaaA	TCAAAACCCT	TGCCGAAGAA	GCCGAACACG	CCAAACAGTC
351	GCTGTCGCTG	ΤΤΤΟΔΔΔΔΤΤ	TCCCCTTCCC	CERTACACG	CCCGTGGACG
401	C1C1C11CC	227 27 27 27 17 1	1000011000	CTACACCGGC	CCCGTGGACG
	GACACAACGT	CGAGAATCTG	GTGGACGTAT	TGAAAGACTT	GCGCAGCCGC
451	AAAGGCCCTC	AGTTGCTGCA	CGTCATCACC	AAAAAGGGCA	ACCCCTACAA
501	ACTCGCCGAA	AACGACCCca	t c = A ATTA CCA	CCCCCCCC	ACGGCIACAA
551	770770000	account	CCAAATACCA	CGCCGTCGCC	aACCTGCcta
	AAGAAGGCGG	GGCGCAAATg	CCGTCTGAAA	AAGAACCCAA	GCCCGCCaca
601	aaaccgaccT	ATACCCAAGT	ATTCGGCAAA	TECCTETECE	*CCCCCCCCC
651	GGCAGATTCC	CGACTCCTTC	CCAmmagaga	2000101000	MCCGGGGGGG
		CONCIDUITO	CGATTACCCC	CGCCATGCGC	CACCCCACCC

7.

. . . . . . . . . . . . . . . .



## This corresponds to the amino acid sequence <SEQ ID 116; ORF 033.ng>: g033.pep

1 MAAADKLLGG DRRSVAIIGD GAMTAGQAFE ALNCAGDMDV DLLVVLNDNE
51 MSISPNVGAL PKYLASNVVR DMHGLLSTVK AQTGKVLDKI PGAMEFAQKV
101 EHKIKTLAEE AEHAKQSLSL FENFGFRYTG PVDGHNVENL VDVLKDLRSR
151 KGPQLLHVIT KKGNGYKLAE NDPVKYHAVA NLPKEGGAQM PSEKEPKPAA
201 KPTYTQVFGK WLCDRAAADS RLVAITPAMR EGSGLVEFEQ RFPDRYFDVG
251 IAEQHAVTFA GGLACEGMKP VVAIYSTFLQ RAYDQLVHDI ALQNLPVLFA
301 VDRAGIVGAD GPTHAGLYDL SFLRCVPNMI VAAPSDENEC RLLLSTCYQA
351 DAPAAVRYPR GTGTGAPVSD GMETVEIGKG IIRREGEKTA FIAFGSMVAT
401 ALAVAEKLNA TVADMRFVKP IDEELIVRLA RSHDRIVTLE ENAEQGGAGG
451 AVLEVLAKHG ICKPVLLLGV ADTVTEHGDP KKLLDDLGLS AEAVERRVRE

## The following partial DNA sequence was identified in N. meningitidis <SEQ ID 117>: m033.seq

ATGGCGGCGG CAGACAAACT CTTGGGCAGC GACCGCCGCA GCGTCGCCAT 1 CATCGGCGAC GGCGCGATGA CGGCGGGGCA GGCGTTTGAA GCCTTGAATT 51 GCGCaG.CGA TATGGATGTr GATTTGCTrG TCGTCCTCAA CGACAACGAA 101 151 ATGTCGATTT CCCCCAACGT CGGCGCGCTG CCGAAATACC TTGCCAGCAA 201 CGTCGTGCGC GATATGCACG GCCTGTTGAG TACCGTCAAA GCGCAAACGG GCAAGGTATT AGACAAAATA CCCGGCGCGA TGGAGTTTGC CCAAAAAGTC GAACACAAA TCAAAACCCT TGCCGAAGAA GCCGAACACG CCAAACAGTC 351 GCTGTCTTG TTTGAAAACT TCGGCTTCCG CTACACCGGC CCCGTGGACG GACACAACGT CGAAAATCTG GTGGACGTAT TGAAAGACTT GCGCAGCCGC AAAGGCCCTC AGTTGCTGCA CGTCATCACC AAAAAGGGCA ACGGCTACAA 451 ACTCGCCGAA AACGACCCCG TCAAATACCA CGCCGTCGCC AACCTGCCTA AAGAAAGCGC GGCGCAAATG CCGTCTGAAA AAGAACCCAA GCCCGCCGCC AAACCGACCT ATACCCAAGT GTTCGGCAAA TGGCTGTGCG ACCGGGCGGC GGCAGATTCC CGACTGGTTG CGATTACCCC CGCCATGCGC GAGGGCAGCG 701 GCTTGGTTGA GTTTGAACAA CGATTCCCCG ACCGCTATTT CGATGTCGGC ATCGCCGAGC AGCACGCCGT TACCTTTGCC GGCGGTTTGG CTTGCGAAGG GATGAAGCCC GTCGTGGCGA TTTATTCCAC CTTTTTACAA CGCGCCTACG 851 ACCAACTGGT GCACGACATC GCCCTGCAAA ACCTACCCGT TTTGTTTGCC GTCGACCGCG CGGGCATCGT CGGCGCGGAC GGCCCGACCC ATGCCGGTCT 951 GTACGATTTG AGCTTTTTGC GCTGCGTGCC GAACATGATT GTCGCCGCGC 1001 CGAGCGATGA AAACGAATGC CGCCTGTTGC TTTCGACCTG CTATCAGGCA 1051 GACGCGCCCG CCGCCGTCCG CTATCCGCGC GGCACGGGTA CGGGCGCGCC 1101 GGTTTCAGAC GGCATGGAAA CCGTGGAAAT CGGCAAGGGC ATTATCCGCC GCGAAGGTGA GAAAACCGCA TTCATTGCCT TCGGCAGTAT GGTCGCCCCC 1151 GCATTGGCGG TTGCCGAAAA ACTGAACGCC ACCGTCGCCG ATATGCGCTT 1251 CGTCAAACCG ATAGACGAAG AGTTGATTGT CCGCCTTGCC CGAAGCCACG



a033.seq



```
1301 ACCGCATCGT TACCCTTGAA GAAAACGCCG AACAGGGCGG CGCAGGCGGC
1351 GCGGTGCTGG AAGTATTGGC GAAACACGGC ATCTGCAAAC CCGTTTTGCT
1401 TTTGGGCGTT GCCGATACCG TAACCGGACA CGGCGATCCG AAAAAACTTT
1451 TAGACGATTT GGGCTTGAGT GCCGAAGCGG TGGAACGGCG TGTGCGCGCG
1501 TGGCTGTCGG ATCGGGATGC GGCAAATTAA
```

## This corresponds to the amino acid sequence <SEQ ID 118; ORF 033>: m033.pep

```
1 MAAADKLLGS DRRSVAIIGD GAMTAGQAFE ALNCAXDMDV DLLVVLNDNE
51 MSISPNVGAL PKYLASNVVR DMHGLLSTVK AQTGKVLDKI PGAMEFAQKV
101 EHKIKTLAEE AEHAKQSLSL FENFGFRYTG PVDGHNVENL VDVLKDLRSR
151 KGPQLLHVIT KKGNGYKLAE NDPVKYHAVA NLPKESAAQM PSEKEPKPAA
201 KPTYTQVFGK WLCDRAAADS RLVAITPAMR EGSGLVEFEQ RFPDRYFDVG
251 IAEQHAVTFA GGLACEGMKP VVAIYSTFLQ RAYDQLVHDI ALQNLPVLFA
301 VDRAGIVGAD GPTHAGLYDL SFLRCVPNMI VAAPSDENEC RLLLSTCYQA
351 DAPAAVRYPR GTGTGAPVSD GMETVEIGKG IIRREGEKTA FIAFGSMVAP
401 ALAVAEKLNA TVADMRFVKP IDEELIVRLA RSHDRIVTLE ENAEQGGAGG
451 AVLEVLAKHG ICKPVLLLGV ADTVTGHGDP KKLLDDLGLS AEAVERRVRA
```

### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 119>:

```
ATGGCGGCGG CGGACAAACA GTTGGGCAGC GACCGCCGCA GCGTCGCCAT
  51
      CATCGGCGAC GGCGCGATGA CGGCGGGTCA GGCGTTTGAA GCCTTGAACT
 101
      GCGCGGCGA TATGGATGTG GATTTGCTGG TCGTCCTCAA CGACAACGAA
     ATGTCGATTT CCCCCAACGT CGGTGCGTTG CCCAAATACC TTGCCAGCAA
      CGTCGTGCGC GATATGCACG GACTGTTGAG TACCGTCAAA GCGCAAACGG
     GCAAGGTATT AGACAAAATA CCCGGCGCGA TGGAGTTTGC CCAAAAAGTC
 251
     GAACATAAAA TCAAAACCCT TGCCGAAGAA GCCGAACACG CCAAACAGTC
 351 ACTGTCTTTG TTTGAAAACT TCGGCTTCCG CTATACCGGC CCCGTGGACG
 401
     GACACAACGT CGAAAATCTG GTCGATGTAT TGGAAGACCT GCGCGGACGC
 451
     AAAGGCCCGC AGCTTCTGCA CGTCATCACC AAAAAGGGCA ACGGCTACAA
501 ACTCGCCGAA AACGATCCCG TCAAATACCA CGCCGTCGCC AACCTGCCTA
 551 AAGAAAGCGC GGCGCAAATG CCGTCTGAAA AAGAACCCAA GCCCGCCGCC
 601 AAACCGACCT ATACCCAAGT GTTCGGCAAA TGGCTGTGCG ACCGGGCGGC
     GGCAGATTCC CGACTGGTTG CGATTACCCC CGCCATGCGC GAGGGCAGCG
     GCTTGGTTGA GTTTGAACAA CGATTCCCCG ACCGCTATTT CGATGTCGGC
 751 ATCGCCGAGC AGCACGCCGT TACCTTTGCC GGCGGTTTGG CTTGCGAAGG
 801 GATGAAGCCC GTCGTGGCGA TTTATTCCAC CTTTTTACAA CGCGCCTACG
 851 ACCAACTGGT GCACGACATC GCCCTGCAAA ACCTGCCCGT TTTGTTTGCC
 901
     GTCGACCGCG CGGGCATCGT CGGCGCGGAC GGCCCGACCC ATGCCGGTTT
     GTACGATTTA AGCTTTTTGC GCTGCATTCC GAATATGATT GTCGCCGCGC
     CGAGCGATGA AAATGAATGC CGCCTGCTGC TTTCGACCTG CTATCAGGCA
1001
1051 GACGCGCCG CCGCCGTCCG CTATCCGCGC GGCACGGGTA CGGGCGTGCC
1101 GGTTTCAGAC GGCATGGAAA CCGTGGAAAT CGGCAAGGGC ATTATCCGCC
1151 GCGAAGGTGA GAAAACCGCA TTCATTGCCT TCGGCAGTAT GGTCGCCCCT
1201
     GCATTGGCGG TCGCCGGAAA ACTGAACGCC ACCGTCGCCG ATATGCGCTT
      CGTCAAACCG ATAGACGAAG AGTTGATTGT CCGCCTTGCC CGAAGCCACG
1301 ACCGCATCGT TACCCTTGAA GAAAACGCCG AACAGGGCGG CGCAGGCAGC
     GCGGTGCTGG AAGTGTTGGC GAAACACGGC ATCTGCAAAC CCGTCTTGCT
1351
     TTTGGGCGTT GCCGATACCG TAACCGGACA CGGCGATCCG AAAAAACTTT
     TAGACGATTT GGGCTTGAGT GCCGAAGCGG TGGAACGGCG TGTGCGCGCG
1451
     TGGCTGTCGG ATCGGGATGC GGCAAATTAA
```

### This corresponds to the amino acid sequence <SEQ ID 120; ORF 033.a>:

```
a033.pep

1 MAAADKQLGS DRRSVAIIGD GAMTAGQAFE ALNCAGDMDV DLLVVLNDNE
51 MSISPNVGAL PKYLASNVVR DMHGLLSTVK AQTGKVLDKI PGAMEFAQKV
101 EHKIKTLAEE AEHAKQSLSL FENFGFRYTG PVDGHNVENL VDVLEDLRGR
151 KGPQLLHVIT KKGNGYKLAE NDPVKYHAVA NLPKESAAQM PSEKEPKPAA
201 KPTYTQVFGK WLCDRAAADS RLVAITPAMR EGSGLVEFEQ RFPDRYFDVG
251 IAEQHAVTFA GGLACEGMKP VVAIYSTFLQ RAYDQLVHDI ALQNLPVLFA
```

301	VDRAGIVGAD	GPTHAGLYDL	SFLRCIPNMI	VAAPSDENEC	RLLLSTCYQA
351	DAPAAVRYPR	GTGTGVPVSD	GMETVEIGKG	IIRREGEKTA	FIAFGSMVAP
401	ALAVAGKLNA	TVADMRFVKP	IDEELIVRLA	RSHDRIVTLE	ENAEQGGAGS
451	AVLEVLAKHG	ICKPVLLLGV	ADTVTGHGDP	KKLLDDLGLS	AEAVERRVRA
501	WLSDRDAAN*				

### m033/a033 98.4% identity over a 509 aa overlap

m033.pep	10 MAAADKLLGSDRRSVA 	20 IIGDGAMTA	30 GQAFEALNCA	40 XDMDVDLLVV	50 LNDNEMSISP	60 NVGAL
a033	MAAADKQLGSDRRSVA 10	IIGDGAMTA	GQAFEALNCA 30	GDMDVDLLVV 40	IIIIIIIIII LNDNEMSISP 50	NVGAL 60
m033.pep	70 PKYLASNVVRDMHGLL	80 STVKAQTGKY	90 VLDKIPGAME:	100 FAQKVEHKIK	110 TLAEEAEHAK	120 QSLSL
a033		 STVKAQTGK 80		 FAQKVEHKIK 100	   LAEEAEHAK   110	IIIII QSLSL 120
m033.pep	130 FENFGFRYTGPVDGHN	140 VENLVDVLKI	150 DLRSRKGPQLI	160 LHVITKKGNG	170 YKLAENDPVK	180 YHAVA
a033		: VENLVDVLE! 140	:       DLRGRKGPQLI 150	 LHVITKKGNG: 160		 YHAVA 180
m033.pep	190 NLPKES <b>AAQMPSEKE</b> PI	200 KPAAKPTYT(	210 QVFGKWLCDRA	220 AAADSRLVAI:	230 FPAMREGSGL	240 VEFEQ
a033					  PAMREGSGL   230	 VEFEQ 240
m033.pep	250 RFPDRYFDVGIAEQHAV	260 /TFAGGLACE	270 EGMKPVVAIYS	280 STFLQRAYDQI	290 LVHDIALQNL	300 PVLFA
a033					 LVHDIALQNL: 290	IIIII PVLFA 300
m033.pep .	310 VDRAGIVGADGPTHAGI	320 LYDLSFLRC\	330 /PNMIVAAPSI	340 DENECRLLLS	350 CYQADAPAA	360 VRYPR
a033		 LYDLSFLRC] 320	:        !  PNMIVAAPSI   330	DENECRLLLST 340	  CYQADAPAA   350	 VRYPR 360
m033.pep	370 GTGTGAPVSDGMETVE	380 IGKGIIRREC	390 SEKTAFIAFGS	400 SMVAPALAVAS	410 EKLNATVADMI	420 RFVKP
a033	:		HIIIIIIIII SEKTAFIAFGS 390			 RFVKP   420
m033.pep	430 IDEELIVRLARSHDRIV	440 /TLEENAEQO	450 GAGGAVLEVI	460 AKHGICKPVI	470 LLLGVADTVT	480 GHGDP
a033			:	 AKHGICKPVI 460	LLLGVADTVTO 470	 GHGDP 480
m033.pep	490 KKLLDDLGLSAEAVERF					
a033						

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae

ORF 033 shows 98.4% identity over a 509 aa overlap with a predicted ORF (ORF 033.ng) from N. gonorrhoeae:

m033/g033

m033.pep	MAAADKLLGSDRRSVAIIGDGAMTAGQAFEALNCAXDMDVDLLVVLNDNEMSISPNVGAL	60
g033	:	60
m033.pep	PKYLASNVVRDMHGLLSTVKAQTGKVLDKIPGAMEFAQKVEHKIKTLAEEAEHAKQSLSL	120
g033	PKYLASNVVRDMHGLLSTVKAQTGKVLDKIPGAMEFAQKVEHKIKTLAEEAEHAKQSLSL	120
m033.pep	FENFGFRYTGPVDGHNVENLVDVLKDLRSRKGPQLLHVITKKGNGYKLAENDPVKYHAVA	180
g033	FENFGFRYTGPVDGHNVENLVDVLKDLRSRKGPQLLHVITKKGNGYKLAENDPVKYHAVA	180
m033.pep	NLPKESAAQMPSEKEPKPAAKPTYTQVFGKWLCDRAAADSRLVAITPAMREGSGLVEFEQ	240
g033	NLPKEGGAQMPSEKEPKPAAKPTYTQVFGKWLCDRAAADSRLVAITPAMREGSGLVEFEQ	240
m033.pep	RFPDRYFDVGIAEQHAVTFAGGLACEGMKPVVAIYSTFLQRAYDQLVHDIALQNLPVLFA	300
g033	RFPDRYFDVGIAEQHAVTFAGGLACEGMKPVVAIYSTFLQRAYDQLVHDIALQNLPVLFA	300
m033.pep	VDRAGIVGADGPTHAGLYDLSFLRCVPNMIVAAPSDENECRLLLSTCYQADAPAAVRYPR	360
g033	VDRAGIVGADGPTHAGLYDLSFLRCVPNMIVAAPSDENECRLLLSTCYQADAPAAVRYPR	360
m033.pep	GTGTGAPVSDGMETVEIGKGIIRREGEKTAFIAFGSMVAPALAVAEKLNATVADMRFVKP	420
g033	GTGTGAPVSDGMETVEIGKGIIRREGEKTAFIAFGSMVATALAVAEKLNATVADMRFVKP	420
m033.pep	IDEELIVRLARSHDRIVTLEENAEQGGAGGAVLEVLAKHGICKPVLLLGVADTVTGHGDP	480
g033		480
m033.pep	KKLLDDLGLSAEAVERRVRAWLSDRDAANX 510	
g033		

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 121>: g034.seq

1	ATGAGCCGTT	TATGGTTTTT	TGCCGTAAAA	AACATTATAA	TCCGCCTTAT
51	TTACCTATTG	CCCAAGGAGA	CACAAATGGC	ACTCGTATCC	ATGCGCCAAC
101	TGCTTGACCA	CGCCGCCGAA	AACAGCTACG	GCCTCCCCC	GTTCAACGTC
151	AACAACCTCG	AACAAATGCG	CCCCATTATC	Charges	ACCAAGTCAA
201	CGCGCCCGTC	ATCCTACACC	CGCCATTATG	GAAGCCGCCG	ACCAAGTCAA
251	CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	AICGIACAGG	CGAGCGCAGG	TGCGCGCAAA	TACGCGGGCG
	CGCCGTTTTT	GCGCCACCTG	ATTCTGGCGG	CAGTCGAAGA	ATTTCCGCAC
301	ATCCCCGTCG	TGATGCACCA	AGACCACGGC	GCATCGCCCG	ACGTALACCA
351	ACGCTCCATC	CAACTGGGCT	TCTCCTCCGT	GATGATGGAC	CCCTCTTTTCC
401	TCGAAGACGG	CAAAACCCCT	TCTTCTTACG	AATACAACGT	CAACCCCAACC
451	CGTACCGTCG	TCAACTTCTC	CCACGCCTGC	CCCCECECEC	TARCUCCACC
501	AATCGGCGTA	TTGGGCAACC	TCCAAACCCC	GGCGTGTCCG	TCGAAGGCGA
551	GAGTGGGGGG	CCCACCCAAA	TCGAAACCGG	CGAAGCAGGC	GAAGAAGACG
601	GAGTGGGCGC	GGCAGGCAAA	CTCTCACACG	ACCAAATGCT	CACCAGCGTT
	GAAGATGCCG	TGCGTTTCGT	TAAAGATACC	GGCGTTGACG	CATTGGCGAT
651	TGCCGTCGGC	ACCAGCCACG	GCGCATACAA	ATTCACCCGT	CCGCCCACAG

ENG .

```
701 GCGACGTATT GCGTATCGAC CGCATCAAGG AAATCCACCA AGCCCTGCCC
751 AATACACACA TCGTGATGCA CGGCCCAGC TCCGTTCCGC AAGAALGGCT
801 GAAAGTCATC AACGAATACG GCGGCAATAT CGGCGAAACC TACGGCGTGC
851 CGGTTGAAGA AATCGTCGAA GGCATCAAAC ACGGCGTGCG CAAAGTCAAC
901 ATCGATACCG ACCTGCGCCT CGCTTCCACC GGCGCGGTAC GCCGCTACCT
951 TGCCGAAAAC CCGTCCGACT TTGATCCGCG CAAATACTTG GGCAAAACCA
1001 TTGAAGCGAT GAAGCAAATC TGCCTCGACC GTTATCTTGC GTTCGGTTGC
1051 GAAGGTCAGG CAGGCAAAAT CAAACCTGTT TCGTTGGAAA AAATGGCAAG
1101 CCGTTATGCC AAGGGCGAAT TGAACCAAAT CGTCAAATAA
```

### This corresponds to the amino acid sequence <SEQ ID 122; ORF 034.ng>: g034.pep

```
1 MSRLWFFAVK NIIIRLIYLL PKETQMALVS MRQLLDHAAE NSYGLPAFNV
51 NNLEQMRAIM EAADQVNAPV IVQASAGARK YAGAPFLRHL ILAAVEEFPH
101 IPVVMHQDHG ASPDVCQRSI QLGFSSVMMD GSLLEDGKTP SSYEYNVNAT
151 RTVVNFSHAC GVSVEGEIGV LGNLETGEAG EEDGVGAAGK LSHDQMLTSV
201 EDAVRFVKDT GVDALAIAVG TSHGAYKFTR PPTGDVLRID RIKEIHQALP
251 NTHIVMHGSS SVPQEWLKVI NEYGGNIGET YGVPVEEIVE GIKHGVRKVN
301 IDTDLRLAST GAVRRYLAEN PSDFDPRKYL GKTIEAMKQI CLDRYLAFGC
351 EGQAGKIKPV SLEKMASRYA KGELNQIVK*
```

## The following partial DNA sequence was identified in N. meningitidis <SEQ ID 123>: m034.seq (partial)

```
1 ATGAGCTGTT TATGGTTTTT TGCTGTAAAA AACATTATAA TCCGCCTTAT
51 TTACCTATTG CCCAAGGAGA CACAAATGGC ACTCGTATCC ATGCGCCAAC
101 TGCTTGATCA TGCTGCCGAA WACAGCTACG GCYTGCCGGC GTTCAACGTC
151 AACAACCTCG WACAGATGCG CGCCATCATG GAGGCTGCAG ACCAAGTCGA
201 CGCCCCCGTC ATCGTACAGG CGAGTGCCGG TGCGCGCAAA TATGCGGGTG
251 CGCCGTTTTT ACGCCACCTG ATTTTGGCGG CTGTCGAAGT ATTTCCACAC
301 ATCCCCGTCG TCATGCACCA AGACCACGGC GCATCACCCG ACGTGTGCCA
351 ACGCTCCATC CAACTGGGCT TCTCCTCTGT AATGATGGAC GGCTCGCTGA
401 TGGAAGACGG CAAAACCCCT TCTTCTTACG AATACAACGT CAACGCCACA
451 CGTACCGTGG TTAACTTCTC CCACGCTTGC GGCGTATCCG TTGAAGGCGA
501 AATCGGCGTA TTGGGCAACC TCGAAACCGG CGATGCAGG GAAGAAGACG
501 GAAGATGCCG TATGTTTCGT TAAAGATACC GGCGTTGACG CATTGGCTAT
601 GAAGATGCCG ACCAGCCACG GCGCATACAA ATTCACCCGT CCGCCCACAG
701 GCGATGTATT ACGTATCGAC CGCCATCAAAG AAATCCACCA AGCCCTGCCC
751 AATACACACA TCGTGATGCA C...
```

### This corresponds to the amino acid sequence <SEQ ID 124; ORF 034>: m034.pep (partial)

```
pep(partial)MSCLWFFAVKNIIIRLIYLLPKETQMALVSMRQLLDHAAEXSYGLPAFNV51NNLXQMRAIMEAADQVDAPVIVQASAGARKYAGAPFLRHLILAAVEVFPH101IPVVMHQDHGASPDVCQRSIQLGFSSVMMDGSLMEDGKTPSSYEYNVNAT151RTVVNFSHACGVSVEGEIGVLGNLETGDAGEEDGVGAVGKLSHDQMLTSV201EDAVCFVKDTGVDALAIAVGTSHGAYKFTRPPTGDVLRIDRIKEIHQALP251NTHIVMH...
```

### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 125>: a034.seq

```
1 ATGAGCCGTT TATGGTTTTT TGCCGCAAAA AACATTATAA TCCGCCTTAT
51 TTACCTATTG CCCAAGGAGA CACAAATGGC ACTCGTATC ATGCGCCAAC
101 TGCTTGATCA TGCTGCCGAA AACAGCTACG GCCTGCCCGC GTTCAACGTC
151 AACAACCTCG AACAAATGCG CGCCATTATG GAAGCCGCCG ACCAAGTCAA
201 CGCGCCGTC ATCGTACAGG CGAGCGCAGG TGCGCGCAAA TACGGGGCG
251 CGCCGTTTTT GCGCCACCTG ATTTTGGCGG CTGTCGAAGA ATTTCCGCAC
301 ATCCCCGTCG TGATGCACCA AGACCACGGC GCATCGCCCG ACGTGTGCCA
351 ACGCTCCATC CAACTGGGCT TTTCCTCCGT GATGATGAC GGCTCGCTGA
401 TGGAAGACGG CAAAACCCCT TCTTCTTATG AATACAACGT CAACGCCACC
```

BNSDCID: -WO__9957280A2_L>

451	CGTACCGTGG	TTAATTTCTC	CCACGCCTGC	GGCGTATCCG	TTGAAGGCGA
501	AATCGGCGTA	TTGGGCAACC	TCGAAACTGG	CGAAGCCGGC	GAAGAAGACC
551	GTGTAGGCGC	AGTGGGCAAA	CTTTCCCACG	ACCAAATGCT	CACCAGCGTC
601	GAAGATGCCG	TGCGTTTCGT	TAAAGATACC	GGCGTTGACG	CATTGGCGAT
651	TGCCGTCGGC	ACCAGCCACG	GCGCGTACAA	ATTCACCCGT	CCCCCCACAC
701	GCGACGTGTT	GCGTATCGAC	CGCATCAAAG	AAATCCACCA	AGCCCTCCCC
751	AATACACACA	TCGTGATGCA	CGGCTCCAGC	TCCGTTCCGC	AAGAATGGCT
801	GAAAGTCATC	AACGAATACG	GCGGCAATAT	CGGCGAAACC	TACGGCGTGC
851	CGGTTGAAGA	AATCGTCGAA	GGCATCAAAC	ACGGCGTGCG	TAAAGTCAAC
901	ATCGATACCG	ACTTGCGCCT	TGCTTCCACC	GGCGCGGTAC	GCCGCTACCT
951	TGCCGAAAAC	CCGTCCGACT	TCGATCCGCG	CAAATATTTG	AGCAAAACCA
1001	TTGAAGCGAT	GAAGCAAATC	TGCCTCGACC	GCTACCTCGC	GTTCGGTTGC
1051	GAAGGTCAGG	CAGGCAAAAT	CAAACCGGTT	TCCTTGGAAA	AAATGGCAAA
1101	CCGTTATGCC	AAGGGCGAAT	TGAACCAAAT	CGTCAAATAA	

### This corresponds to the amino acid sequence <SEQ ID 126; ORF 034.a>:

### m034/a034 96.9% identity over a 257 aa overlap

	10	20	30	40	50	60
m034.pep	MSCLWFFAVKNII	IRLIYLLP <b>KE</b> '	TOMALVSMRO	LLDHAAEXSY	GLPA FNUNNI.	MT A DMOY.
			11111111		I I I I I I I I I I I I I I I I I I I	
a034	MSRLWFFAAKNII:	RLIYLLPKE'	COMALVSMRO	LLDHAAENSV	יון ווווון וו	FONDETA
	10	20	30	40		
			30	40	50	60
	70	80	90	100	110	100
m034.pep	EAADQVDAPVIVQA	SAGARKYAGA	APFIRHT.TT.AZ	17517500 T D177	110	120
a034	EAADOVNAPVTVO	SACABEVACI				11111
	EAADQVNAPVIVQA 70	O O O	FEL TEUTITE		VMHQDHGASP:	DVCQRSI
	70	80	90	100	110	120
	100					
-024	130	140	150	160	170	180
m034.pep	QLGFSSVMMDGSLM	IEDGKTPSSYE	YNVNATRTVV	NFSHACGVS	EGEIGVI.GN	LETCDAC
			1111111			
a034	QLGFSSVMMDGSLM	EDGKTPSSYE	YNVNATRTVV	NESHACGVSV	TEGET GVI CNI	
	130	140	150	160	170	
				100	170	180
	190	200	210	220	220	
m034.pep	EEDGVGAVGKLSHD	OMITSVEDAV		TATALLOMO	230	240
		111111111	CIVEDIGUDA	LAIAVGISHG	AYKITRPPTO	SDVLRID
a034			111111	111111111		
	EEDGVGAVGKLSHD 190	OMP 1 2 A FDWA	REVEDIGVDA		AYKFTRPPTO	SDVLRID
	190	200	210	220	230	240
	25.0					
m034	250					
m034.pep	RIKEIHQALPNTHI					
		111				
a034	RIKEIHQALPNTHI	VMHGSSSVPQ	EWLKVINEYG	GNIGETYCVÞ	VERTURGIVE	Cungan
	250	260	270	280	290	
				200	290	300

Computer analysis of this amino acid sequence gave the following results: Homology with a predicted ORF from N. gonorrhoeae

ORF 034 shows 96.5% identity over a 257 aa overlap with a predicted ORF (ORF 034.ng) from N. gonorrhoeae:

m034/g034		
m034.pep	MSCLWFFAVKNIIIRLIYLLPKETQMALVSMRQLLDHAAEXSYGLPAFNVNNLXQMRAIM	60
g034		60
m034.pep	EAADQVDAPVIVQASAGARKYAGAPFLRHLILAAVEVFPHIPVVMHQDHGASPDVCQRSI	120
g034	EAADQVNAPVIVQASAGARKYAGAPFLRHLILAAVEEFPHIPVVMHQDHGASPDVCQRSI	120
m034.pep	QLGFSSVMMDGSLMEDGKTPSSYEYNVNATRTVVNFSHACGVSVEGEIGVLGNLETGDAG	180
g034	QLGFSSVMMDGSLLEDGKTPSSYEYNVNATRTVVNFSHACGVSVEGEIGVLGNLETGEAG	180
m034.pep	EEDGVGAVGKLSHDQMLTSVEDAVCFVKDTGVDALAIAVGTSHGAYKFTRPPTGDVLRID	240
g034	EEDGVGAAGKLSHDQMLTSVEDAVRFVKDTGVDALAIAVGTSHGAYKFTRPPTGDVLRID	240
m034.pep	RIKEIHQALPNTHIVMH 	257
g034	RIKEIHQALPNTHIVMHGSSSVPQEWLKVINEYGGNIGETYGVPVEEIVEGIKHGVRKVN	300
	partial DNA sequence was identified in N. gonorrhoeae <seq 127="" id="">:</seq>	
g036. <b>se</b> q	AMOCHOLINGO COMOMMUCCH AND CLOSE MOREOGRAPHIC COMO	
1 51	ATGCTGAAGC CGTGTTTGGT ATACAGTGCC TGTGCGGCGG cgttgcCTGC GCGGACTTCG AGCAGCAGGC GTTGCGTGCC TTCGGGCAGA TGTGCGTACC	
101	AATATTCGAG CAGGGCGGAC GCAACGCCCC GTCGGCGGCA TTCGGGCGCG	
151	GTGGCAATCA GGTGCAGTTC GGATTCGTCG GGCAGGTTCT GCCAAACGAT	
201	AAAGGCGGCA ATCCTGCCGT CTTTTTCCGC AAGGAAAACC TGTTCGGACG	
251	GCGAAACAAG CGCGGACTCA AATTGGCGTT GCGTCCACGC GGACGGGTTG	
301	CAGACGGTAT CGAGCGCGGC CAGTGCGGCG CAGTCGGACG GTGAGGCTGG	
351	GCGGATGTTC ATGTTCGTGC CTTCCGTTCC GCCTGTTCTT TGGCAGTCAG	
401	GGCGATTTTG TTGCGGACGT AGAGCAGTTC GGCGTGTGCC GCGCCAGTTG	
451	CGGGATAGCC GCCGCCGAGG GCGAGCGCGA GAAAATCGGC GGCGGTCGGC	
501	ATATCGGGTT TGCCTGAGAA GGGCGGACGG TTTTCCAGTG CGAACGCACT	
551	GCCGATGCCG TCTGAAAAGA CGTACCCCTC GGGGAGGGCA ATGTCTGCCG	
601	CCCTACCGAC TTGATAATCG CTCAAACGGC GGCGGTTCAG CGTGTCGAAC	
651	CACGCATAAA ACACTTCGCC CATACGCGCG TCCGCAGCGG CGAGTATGCA	
701	GCTTTGCGGC GGCGGCAGCG AGGCGGCGGC ATCGAGCGTG GGGATGCCGA	
751	TTAAAGGCGT GTCGAACGGC GTTGCCAAAC CTTGCGCCAC GCCGATGCCG	
801	ATACGCAGTC CGGTAA	
_	s to the amino acid sequence <seq 036.ng="" 128;="" id="" orf="">:</seq>	
<b>g036.pep</b>	MLKPCLVYSA CAAALPARTS SSRRCVPSGR CAYQYSSRAD ATPRRRHSGA	
51		
101		
	RDSRRGRAR ENRRRSAYRV CLRRADGFPV RTHCRCRLKR RTPRGGQCLP	
201	PYRLDNRSNG GGSACRTTHK TLRPYARPQR RVCSFAAAAA RRRHRAWGCR	
251		
The following p	artial DNA sequence was identified in N. meningitidis <seq 129="" id="">:</seq>	
m036.seq		
1		
51		
101		
151		
201		
251		
301	CAGACGGCAT CGAGTGCGGC CAGCTCCTCA CAATCGGCAC AAACGGCACG	

```
351 GCGGATGTTC ACGGGCGCG TCTCCGTTCG GCCTGTTCTT TGGCAGTCAG
401 GGCGATTTTG TTGCGGACGT AGAGCAAACC GGCGTGTGCG GCATGGACGG
451 CAGGATAACC GCCCTTGGCT GCCAATGCGA GAAAGTCGGC GGCAGTCGGC
501 ATATCCGGTC TGCCTGAGAA CGGCGGAGCT TCTTCCAGCG CGAACGCGCT
551 GCCTATGCCG TCTGAAAAGG CGCATCCCTC CGGCAGCCGG ATGTCTGCCG
601 CCCGCCCGAC CTGATAATCG CTCAAACGGT GGCAGTTCAG CGTATCGAAC
651 CATGCATAAA ACACTTCGCC CATACGAGCG TCCGTAGCGG CAAGGATGCA
701 GCTTTGCGGC GGCGGCAGCG AGGCGGCGC ATCGAGCAG GGTACGCCGA
751 TTAAGGGGGT ATCAAACGGC GTTGCCAAAC CCTGAGCTAC ACCGATGCCG
801 ATACGCAGTC CGGTAA
```

## This corresponds to the amino acid sequence <SEQ ID 130; ORF 036>: m036.pep

```
1 MLKPCAVYSA CAAVLPARTS SSRRCVSSGR CVNQYSSRAD AIPWRRHSGA
51 VAIRCSSDSS GRFCQTIKAA IPXSFSARKT CSDGETSADS NWRCVHADGL
101 QTASSAASSS QSAQTARRMF TGALSVRPVL WQSGRFCCGR RANRRVRHGR
151 QDNRPWLPMR ESRRQSAYPV CLRTAELLPA RTRCLCRLKR RIPPAAGCLP
201 PARPDNRSNG GSSAYRTMHK TLRPYERP*R QGCSFAAAAA RRRHRARVRR
251 LRGYQTALPN PELHRCRYAV R*
```

## The following partial DNA sequence was identified in N. meningitidis <SEQ ID 131>:

J. 364					
1	ATGCTGAAGC	CGTGCGCCGT	GTACAGTGCC	TGTGCGGCGG	TGTTGCCTGC
51	ACGGACTTCG	AGCAGCAGGC	GTTGCGTGTC		TGTGTGAACC
101	AATATTCGAG	CAGGGCGGAC	GCAATTCCTT		TTCGGGCGCG
151	GTGGCAATCA	GGTGCAGTTC	GGATTCGTCG		GCCAAACGAT
201	AAAGGCGGCA	ATCCCGCCGT	CTTTTTCCGC		TGTTCGGACG
251		TGCGGACTCA		GCGTCCACGC	GGACGGGTTG
301		CGAGCGCGGC		CAATCGGCAT	AAACGGCGCG
351		ACAGGCGCGC		GCCTGTTCTT	TGGCAGTCAA
401	GGCGATTTTG	TTGCGGACGT	AGAGCAGCTC	GGCGTGTGCC	GCAGCGACGG
451	CGGGAAAACC	GCCTTCAGCC	GCCAGATTGA	GGAAGTCGGC	GGCGGTCGGC
501	ATATCGGGTT	TGCCTGAGAA	GGGCGGACGG	TTTTCCAGCG	CGAACGCATT
551	GCCGATGCCG	TCTGAAAAGG	CGCATCCTTC	CGGCAGCCGG	ATGTCTGCCG
601	CCCGACCGAC	CTGATAATCG	CTCAAACGGC	GGCGGTTCAG	CGTGTCGAAC
651	CATGCATAAA	ACACTTCGCC	CATACGTGCG	TCCGCAGCGG	CAAGGATGCA
701	GCTTTGCGGC	GGCGGCAGCG	AGGCGGCGGC	ATCGAGCGAG	GGTACGCCGA
751	TTAAAGGAGT	ATCAAACGGC	GTTGCCAAAC	CTTGCGCCAC	GCCGATGCCG
801	ATACGCAGTC	CCGTAA			

### This corresponds to the amino acid sequence <SEQ ID 132; ORF 036.a>:

a036.pep					
1	MLKPCAVYSA	CAAVLPARTS	SSRRCVSSGR	CVNOYSSRAD	ATPWRRHSCA
51	VAIRCSSDSS	GRFCQTIKAA	IPPSFSARKT	CSDGETSADS	NWRCVHADGI
101	QTASSAASAA	QSA*TARRMF	TGAPSVPPVL	WOSRRFCCGR	RAARRVPORR
151	RENRLQPPD*	GSRRRSAYRV	CLRRADGFPA	RTHCRCRLKR	RTIPAACCIP
201	PDRPDNRSNG	GGSACRTMHK	TLRPYVRPOR	OGCSFAAAAA	RESHEARVER
251	LKEYOTALPN	LAPRRCRYAV	D*		

### m036/a036 85.6% identity over a 270 aa overlap

m036.pep	10 MLKPCAVYSACAAV	20 LPARTSSSRF	30 CVSSGRCVNO			
a036	MLKPCAVYSACAAV	LPARTSSSRF	CVSSGRCVNO		ILLIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	
	10	20	.o.roooncvno		KKUSGAVAII	RCSSDSS
	10	20	30	40	50	60
m036.pep	70	80	90	100	110	120
mose.pep	GRFCQTIKAAIPXS	FSARKTCSDG	ETSADSNWRC	VHADGLOTAS	SAASSSOSAC	TARRME
a036		111111	111111111	1111111111	1111111	11111
8030	GRFCQTIKAAIPPS	FSARKTCSDG	ETSADSNWRC	VHADGLQTAS	SAASAAOSAX	TARRME
	70	80	90	100	110	120

m036.pep	130 TGALSVRPVLWQSG					
a036	TGAPSVPPVLWQSR		:  ::		:	: :
2000	130	140	150	160	170	180
	190	200	210	220	230	240
m036.pep	RTRCLCRLKRRIPP	AAGCLPPARF	DNRSNGGSSAY	/RTMHKTLRP	YERPXRQGCS	
	11:1 111111		11111111:11	11111111	1 11 11111	
a036	RTHCRCRLKRRILP	AAGCLPPDRP	DNRSNGGGSA	CRTMHKTLRP	YVRPQRQGCS	FAAAAA
	190	200	210	220	230	240
	250	260	270			
m036.pep	RRRHRARVRRLRGY	<b>TALPNPELH</b>	RCRYAVRX			
	111111111111111111111111111111111111111	: 41111	111111			
a036	RRRHRARVRRLKEY	<b>TALPNLAPR</b>	RCRYAVPX			
	250	260	270			

Computer analysis of this amino acid sequence gave the following results: Homology with a predicted ORF from N. gonorrhoeae

ORF 036 shows 74.9% identity over a 271 aa overlap with a predicted ORF (ORF 036.ng) from N. gonorrhoeae:

m036/g036

	10	20	30	40	50	60
m036.pep	MLKPCAVYSACAAV:	LPARTSSSRR	CVSSGRCVNQ	YSSRADAIP	RRHSGAVAI	RCSSDSS
		1111111111		1111111	111111111	
g036	MLKPCLVYSACAAA		CVPSGRCAYQ	YSSRADATPF	RRHSGAVAIR	RCSSDSS
	10	20	30	40	50	60
006	70	80	90	100	110	120
m036.pep	GRFCOTIKAAIPXS					TARRMF
006						1 111
g036	GRFCQTIKAAILPS					EAGRMF
	70	80	90	100	110	120
	120	1.10				
0.2.6	130	140	150	160	170	180
m036.pep	TGALSVRPVLWQSGI					AELLPA
-026	:		11: :1:1			1: :1:
g036	MFVPSVPPVLWQSGI	RECCGRRAVR	RVPRQLRDSR			RADGFPV
	130	140	150	160	170	180
	190	200	010			
-026		200	210	220	230	240
m036.pep	RTRCLCRLKRRIPPA					FAAAAA
~03 <i>c</i>		:     .	111111:11			11111
g036	RTHCRCRLKRRTPRO	GQCLPPYRL				
	190	200	210	220	230	240
	250	0.60	070			
-036	250	260	270			
m036.pep	RRRHRARVRRLRGY					
~036	::::					
g036	RRRHRAWGCRLKACE					
	250	260	270			

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 133>: m036-1.seq

1	ATGCTGAAGC	CGTGCGCCGT	GTACAGTGCC	TGTGCGGCGG	TGTTGCCTGC
51	ACGGACTTCG	AGCAGCAGGC	GTTGCGTGTC	TTCGGGCAGA	TGTGTGAACC
101	AATATTCGAG	CAGGGCGGAC	GCAATTCCTT	GGCGGCGCA	TTCGGGCGCG
151	GTGGCAATCA	GGTGCAGTTC	GGATTCGTCG	GGCAGGTTCT	GCCAAACGAT
201	AAAGGCGGCA	ATCCCGCCGT	CTTTTTCCGC	AAGGAAAACC	TGTTCGGACG
251	GCGAAACCAG	TGCGGACTCA	AATTGGCGTT	GCGTCCATGC	GGACGGGTTG
301	CAGACGGCAT	CGAGTGCGGC	CAGCTCCTCA	CAATCGGCAC	AAACGGCACG

BNSDOCID: <WO___8957280A2_I_>

351 401 451 501 551	GGCGATTTTG CAGGATAACC ATATCCGGTC GCCTATGCCG	TTGCGGACGT GCCCTTGGCT TGCCTGAGAA TCTGAAAAGG	AGAGCAAACC GCCAATGCGA CGGCGGAGCT CGCATCCCTC	GGCGTGTGCG GAAAGTCGGC TCTTCCAGCG CGGCAGCCGG	ATCTCTCCCC
<b>6</b> 01	CCCGCCCGAC CATGCATAAA	CTGATAATCG	CTCAAACGGT	GGCAGTTCAG	ATGTCTGCCG CGTATCGAAC

#### This corresponds to the amino acid sequence <SEQ ID 134; ORF 0036-1>: m036-1.pep

- 1 MLKPCAVYSA CAAVLPARTS SSRRCVSSGR CVNQYSSRAD AIPWRRHSGA 51
- VAIRCSSDSS GRFCQTIKAA IPPSFSARKT CSDGETSADS NWRCVHADGL
- QTASSAASSS QSAQTARRMF TGALSVRPVL WQSGRFCCGR RANRRVRHGR 151
- QDNRPWLPMR ESRRQSAYPV CLRTAELLPA RTRCLCRLKR RIPPAAGCLP
- 201 PARPDNRSNG GSSAYRTMHK TLRPYERP*

#### 76.8% identity in 228 aa overlap m036-1/g036

```
30
                                              40
            MLKPCAVYSACAAVLPARTSSSRRCVSSGRCVNQYSSRADAIPWRRHSGAVAIRCSSDSS
m036-1.pep
            MLKPCLVYSACAAALPARTSSSRRCVPSGRCAYQYSSRADATPRRRHSGAVAIRCSSDSS
g036
                   10
                            20
                                     30
                                              40
                                                       50
                   70
                            80
                                     90.
                                             100
                                                      110
                                                               120
            GRFCQTIKAAIPPSFSARKTCSDGETSADSNWRCVHADGLQTASSAASSSQSAQTARRMF
m036-1.pep
            {{{\bar{1}}}}
g036
            GRFCQTIKAAILPSFSARKTCSDGETSADSNWRCVHADGLQTVSSAASAAQSDGEAGRMF
                   70
                            80
                                     90
                                             100
                                                      110
                                                               120
                           140
                                    150
                                             160
           TGALSVRPVLWQSGRFCCGRRANRRVRHGRQDNRPWLPMRESRRQSAYPVCLRTAELLPA
                                                      170
m036-1.pep
              : 11 11111111111111 111 : :1:1
                                              {\tt MFVPSVPPVLWQSGRFCCGRRAVRRVPRQLRDSRRRGRARENRRRSAYRVCLRRADGFPV}
g036
                  130
                           140
                                    150
                                             160
                                                               180
                  190
                           200
                                    210
                                             220
m036-1.pep
           RTRCLCRLKRRIPPAAGCLPPARPDNRSNGGSSAYRTMHKTLRPYERPX
           ##:| ##|#|# | :: #|#| | ##|##|:|| ## ||##|## #|
           RTHCRCRLKRRTPRGGQCLPPYRLDNRSNGGGSACRTTHKTLRPYARPQRRVCSFAAAAA
g036
                           200
                                    210
                                             220
                                                      230
g036
           RRRHRAWGCRLKACRTALPNLAPRRCRYAVRX
                  250
                                    270
```

#### The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 135>: g038.seq

1	ATGACTGATT	TCCGCCAAGA	TTTCCTCAAA	TTCTCCCTCG	СССААААТСТ
51	TTTGAAATTC	GGCGAATTTA	CCACCAAAGC	CGGACGCCG	TCCCCCmamm
101	TCTTCAATGC	CGGCCTCTTC	AACGACGGCG	CGTCCACGCT	CCAACTCCCTATT
151	AAATTCTATG	CACAATCCAT	CATTGAAAGC	CCCATCCCAT	GCAACTGGCA
201	GTTCGGCCCC	GCCTACAAAG	GCATTATTTT	GCCCCCCCAT	TCGATATGCT
251	TGCTGGCGGA	AAAAGGCGTG	AACGTCCCGT	TTTCCCTT CAR	ACCGCGATGA
301	GCCAAAGACC	GCGGCGAAGG	CGGCGTGTTG	CTCCCTACAA	CCGCAAAGAA
351	GCGCGTGCTG	ATTATCGACG	ACGTGATTTC	GICGGGGG	CGCTTAAAGG
401	AATCAATCAA	ACTCATOCAA	CCCCACCCC	CGCCGGCACA	TCCGTACGCG
451	ATCGCGCTCG	VCCCCV MCCV	GCGGAGGGTG	CAACCCCCGC	CGGTGTCGCC
501	ATCGCGCTCG	ACCGCATGGA	AAAAGGCACG	GGTAAATTGT	CCGCCGTTCA
551	GGAAGTGGAA	MANCARTACG	GCCTGCCCGT	CGCCCCCATC	GCCAGCCTGA
601	ACGATTTGTT	TATCCTGTTG	CAAAACAACC	CCGAATTCGG	ACAGTTCCTC
001	GAACCCGTCC	GCACCTACCG	CCGGCAGTAC	GGCGTAGAAT	AA

```
This corresponds to the amino acid sequence <SEQ ID 136; ORF 038.ng>:
      g038.pep
                MTDFRQDFLK FSLAQNVLKF GEFTTKAGRR SPYFFNAGLF NDGASTLQLA
            51
               KFYAQSIIES GIRFDMLFGP AYKGIILAAA TAMMLAEKGV NVPFAYNRKE
           101 AKDRGEGGVL VGAPLKGRVL IIDDVISAGT SVRESIKLIE AEGATPAGVA
           151 IALDRMEKGT GKLSAVQEVE KQYGLPVAPI ASLNDLFILL QNNPEFGQFL
           201 EPVRTYRRQY GVE*
The following partial DNA sequence was identified in N. meningitidis <SEO ID 137>:
     m038.seq
                ATGACCGATT TCCGCCAAGA TTTCCTCAAA TTCTCCCTCG CCCAAAATGT
             1
            51
                TTTGAAATTC GGCGAATTTA CCACCAAGGC AGGACGGCGG TCGCCCTATT
               TCTTCAATGC CGGCCTCTTT AACGACGGCT TGTCCACGCT GCAACTGGCA
           101
           151 AAATTTTACG CACAATCCAT CATTGAAAGC GGCATCCGAT TCGATATGCT
           201 GTTCGGTCCC GCCTACAAAG GCATTATTTT GGCGGCGGCA ACCGCGATGA
           251 TGCTGGCGA AAAAGGCGTG AACGTCCCGT TTGCCTACAA CCGCAAAGAA
               GCCAAAGACC ACGGCGAAGG CGGCGTGTTG GTCGGCGCGC CGCTTAAAGG
           351 GCGCGTGCTG ATTATCGACG ACGTGATTTC CGCCGGCACA TCCGTACGCG
           401 AATCGATCAA ACTGATTGAA GCGGAGGGTG CAACCCCcGC CGGTGTCGCC
           451 ATCGCGCTCG ATCGCATGGA AAAAGGCACG GGTGAATTGA GCGCGGTTCA
           501 GGAAGTGGA: AAACAATACG GKCTGCCCGT CGCCCCCATC GCCAGCCTGA
               ACGATTTGTT TATTCTGTTG CAAAACAACC CCGAATTCGG ACAGTTCCTC
               GAACCCGTCC GAGCCTACCG TCGGCAGTAC GGCGTAGAAT AA
This corresponds to the amino acid sequence <SEQ ID 138; ORF 038>:
     m038.pep
            1 MTDFRQDFLK FSLAQNVLKF GEFTTKAGRR SPYFFNAGLF NDGLSTLOLA
           51 KFYAQSIIES GIRFDMLFGP AYKGIILAAA TAMMLAEKGV NVPFAYNRKE
           101
               AKDHGEGGVL VGAPLKGRVL IIDDVISAGT SVRESIKLIE AEGATPAGVA
               IALDRMEKGT GELSAVQEVE KQYGLPVAPI ASLNDLFILL QNNPEFGQFL
          201 EPVRAYRRQY GVE*
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 139>:
     a038.seq
               ATGACCGATT TCCGCCAAGA TTTCCTCAAA TTCTCCCTCG CCCAAAATGT
            1
           51
               TTTGAAATTC GGCGAATTCA CCACCAAAGC CGGACGGCGG TCGCCCTATT
               TCTTCAATGC CGGCCTCTTT AACGACGGCT TGTCCACGCT GCAACTGGCA
          101
          151 AAATTTTACG CACAATCCAT CATTGAAAGC GGCATCCGAT TCGATATGCT
          201 GTTCGGCCCC GCCTACAAAG GCATTATTTT GGCGGCGGCA ACCGCGATGA
          251 TGCTGGCGGA AAAAGGCGTG AACGTCCCGT TTGCCTACAA CCGCAAAGAA
          301 GCCAAAGACC ACGGCGAAGG CGGCGTGTTG GTCGGCGCGC CGCTTAAAGG
          351 GCGCGTGCTG ATTATCGACG ACGTGATTTC CGCCGGCACA TCCGTACGCG
401 AATCGATCAA ACTGATTGAA GCGGAGGGTG CAACCCCCGC CGGTGTCGCC
          451 ATCGCGCTCG ACCGCATGGA AAAAGGCACG GGTGAATTGA GCGCGGTTCA
          501 GGAAGTGGAA AAACAATACG GCCTGCCCGT CGCCCCCATC GCCAGCCTGA
          551 ACGATTTGTT TATTCTGTTG CAAAACAACC CCGAATTCGG ACAGTTCCTC
          601 GAACCCGTCC GAGCCTACCG TCGGCAGTAC GGCGTAGAAT AA
This corresponds to the amino acid sequence <SEQ ID 140; ORF 038.a>:
     a038.pep
               MTDFRQDFLK FSLAQNVLKF GEFTTKAGRR SPYFFNAGLF NDGLSTLQLA
           51 KFYAQSIIES GIRFDMLFGP AYKGIILAAA TAMMLAEKGV NVPFAYNRKE
          101 AKDHGEGGVL VGAPLKGRVL IIDDVISAGT SVRESIKLIE AEGATPAGVA
          151 IALDRMEKGT GELSAVQEVE KQYGLPVAPI ASLNDLFILL QNNPEFGQFL
          201 EPVRAYRRQY GVE*
             100.0% identity over a 213 aa overlap
m038/a038
                           10
                                     20
                                               30
                                                          40
                  MTDFRQDFLKFSLAQNVLKFGEFTTKAGRRSPYFFNAGLFNDGLSTLQLAKFYAQSIIES
```

BNSDOCID: -WO__9957280A2_J_>

a038		AQNVLKFGEFT 20	TKAGRRSPYI 30	 FFNAGLFNDG] 40	 LSTLQLAKFY 50	 AQSIIES 60
m038.pep	70 GIRFDMLFGPAYK           GIRFDMLFGPAYK 70		  LAEKGVNVPE	 AYNRKEAKDI		 PLKGRVL
m038.pep	130 IIDDVISAGTSVRI	140	90 150	100 160	110 170	120 180
a038		11111111	11111111		111111111	111111
m038.pep	190 ASLNDLFILLQNNI            ASLNDLFILLQNNI 190		111111111			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae

ORF 038 shows 98.1% identity over a 213 aa overlap with a predicted ORF (ORF 038.ng) from N. gonorrhoeae:
m038/g038

222	10	20	30	40	50	60
m038.pep	MTDFRQDFLKFSLA	AQNVLKFGEF!	TTKAGRRSPYE	FNAGLFNDGL	STLQLAKFY	AOSIIES
222					111111111	111111
g038	MTDFRQDFLKFSLA	AQNVLKFGEF1	TKAGRRSPYF	FNAGLFNDGA	STLOLAKFY	AOSTTES
	10	20	30	40	50	60
						•
	70	80	90	100	110	120
m038.pep	GIRFDMLFGPAYK	MATAAAIII	<i>ILAEKGV</i> NVPF	'AYNRKEAKDH	GEGGVLVGA	PLKGRVI.
		#		111111111:	1111111111	LILITIA
g038	GIRFDMLFGPAYKG	IILAAATAMM	ILAEKGVNVPF	AYNRKEAKDR	GEGGVLVGA	PLKGRVI.
	70	80	90	100	110	120
	130	140	150	160	170	180
m038.pep	IIDDVISAGTSVRE	SIKLIEAEGA	TPAGVAIALD	RMEKGTGELS.	AVOEVEKOYO	TPVAPT
		1111111111		1111111:11	THILLIAM	111111
g038	IIDDVISAGTSVRE	SIKLIEAEGA	TPAGVAIALD	RMEKGTGKLS	AVOEVEKOYO	TPVAPT
	130	140	150	160	170	180
						200
	190	200	210			
m038.pep	ASLNDLFILLQNNP	EFGQFLEPVR	AYRRQYGVEX			
	_	1111111111	: ! ! ! ! ! ! ! ! ! !			
g038	ASLNDLFILLQNNP	EFGQFLEPVR	TYRRQYGVEX			
	190	200	210			

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 141>:

1 ATGCCGTCCG AACCACCTGC CGCTTCAGAC GGCATCAAAC CGACACACACACACACACACACACACACACACAC	1	ATGCCGTCCG	AACCACCTGC	CGCTTCAGAC	GGCATCAAAC	CGACACACAC
101 CTGGGTCAAA GAAACCCAGC TCAACGtCGC CCAAGGCTTC GTCGTCTGC	51	CGAGAAAACA	TCATGCCCGC	CTGTTTCTGT	CCGCACTGCA	AAACCCCCCCC
151 aaAAAtgcga agGGCTgttt aaAgccaaaG accAtctggg aacgard	101	CTGGGTCAAA	GAAAcccagC	TCAAcatCaC	CCARCCCTTC	CTCCTCTC
	151	aaAAAtqcqa	agGGCTgttt	33400032030	Scaltatore	Greenergee
201 gaacctatat tcaacgattg gcccqaaqct gtttcgggat gTcaaaCTCc	201	gaacctatat	traargatto	CCCCCCCCCC	accatetgge	aaGCaCGAAA

```
251 TCcaccgcaT cggcacgcac gccattagca aGAaacagat gtcccgcgac
          301 gaaatCgccg atatcctcaa cggcggtaca acCCTGCACG ATACGCCGCC
          351 CGCAACCGCC GCTGCCGCac ctGCCGCCGC ACCGCaqqTT TCCGTACCGC
          401 CCGCCGTCA GGAAGGGCTC AACTGGACTA TTGCAACCCT GTTCGCACTT
          451 ATCGTCCTCA TTATGCAGCT TTCCTACCTC TTCATCCTAT GA
This corresponds to the amino acid sequence <SEQ ID 142; ORF 039.ng>:
     g039.pep
               MPSEPPAASD GIKPTHTEKT SCPPVSVRTA KPASGSKKPS STSPKASSSA
            1
           51
               KNAKGCLKPK TIWQARKNLY STIGPKLFRD VKLVHRIGTH AISKKQMSRD
          101
               EIADILNGGT TLHDTPPATA AAAPAAAPQV SVPPARQEGL NWTIATLFAL
               IVLIMQLSYL FIL*
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 143>:
     m039.seg
               ATGCCGTCCG AACCGCCTTA CGCCTCAGAC GGCATCAAAC CTGACACACA
            1
               CGAGGAAATA CCATGCCCGC CTGTTTCTGC CCCCACTGCA AAACCCGTCT
           51
          101 CTGGGTCAAA GAAACCCAAC TCAATGTCGC CGnnnnnnn nnnnnnnnn
          201 nnnnnnnn nnnnnnnnn nnnnnnnnn nnnnnnccc GAGGCTGTTT
          251 CGGATGTCAA ACTCGTTCAC CGTATCGGCA CGCGCGCCAT CGGCAAGAAA
          301 CAGATTTCCC GTGACGAAAT CGCCGGCATC CTCAACGGCG GTACAACCCA
          351 GCCCGATATT CCGCCCGCAA CCGCCGCCAC CCCTGCTGCC GCACCGCAGG
          401 TTACCGTACC GCCCGCCGCG CCCGCCCGTC AGGATGGGTT CAACTGGACG
          451 ATTGCAACCC TGTTTGCCCT TATCGTCCTC ATTATGCAGC TTTCCTACCT
          501 CGTCATCCTA TGA
This corresponds to the amino acid sequence <SEQ ID 144; ORF 039>:
     m039.pep
              MPSEPPYASD GIKPDTHEEI PCPPVSAPTA KPVSGSKKPN SMSPXXXXXX
           51
             XXXXXXXXX XXXXXXXXX XXXXXXXXP EAVSDVKLVH RIGTRAIGKK
          101
               QISRDEIAGI LNGGTTQPDI PPATAATPAA APQVTVPPAA PARQDGFNWT
          151 IATLFALIVL IMQLSYLVIL *
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 145>:
     a039.seq
               ATGCCGTCTG AACCGCCTTA CGCCTCAGAC GGCATCAAAC CTGACACACA
           1
              CGAGGAAATA CCATGCCCGC CTGTTTCTGC CCCCACTGCA AAACCCGTCT
           51
          101
              CTGGGTCAAA GAAACCCAAC TCAATGTCGC CCAAGGCTTC GTCGTCTGCC
              AAAAATGCGA AGGAATGTTT AAAGCCAAAG ACCATCTGGC AAGCACGAAA
          151
          201 GAACCCATAT TCAACGATT. TGCCCGAAGC TGTTTCGGAT GTCAAACTCG
              TTCACCGCAT CGGCACGAGC GCCATCGGCA AGAAACAGAT TTCCCGTGAC
          251
          301 GAAATCGCCG GCATCCTCAA CGGCGGCACA ACCCAGCCCG ATATTCCGCC
              CGCAACCGCC GCCACCCCTG CTGCCGCACC GCAGGTTACC GTACCGCCCG
          351
               CCGCGCCCGC CCGTCAGGAT GGGTTCAACT GGACGATTGC AACCCTGTTT
          451 GCCCTTATCG TCCTCATTAT GCAGCTTTCC TACCTCGTCA TCCTATGA
This corresponds to the amino acid sequence <SEQ ID 146; ORF 039.a>:
     a039.pep
              MPSEPPYASD GIKPDTHEEI PCPPVSAPTA KPVSGSKKPN SMSPKASSSA
              KNAKECLKPK TIWQARKNPY STIXPEAVSD VKLVHRIGTS AIGKKQISRD
              EIAGILNGGT TOPDIPPATA ATPAAAPQVT VPPAAPARQD GFNWTIATLF
              ALIVLIMOLS YLVIL*
m039/a039
           79.4% identity over a 170 aa overlap
                                   20
                                            30
                                                      40
                                                                50
                 MPSEPPYASDGIKPDTHEEIPCPPVSAPTAKPVSGSKKPNSMSPXXXXXXXXXXXXXXXXXX
     m039.pep
                 a039
                 MPSEPPYASDGIKPDTHEEIPCPPVSAPTAKPVSGSKKPNSMSPKASSSAKNAKECLKPK
```

	10	20	30	40	50	60
m039.pep	70 XXXXXXXXXXXXXX : :   TIWQARKNPYSTIX- 70	 PEAVS	DVKLVHRIGT	 SAIGKKQIS	11111111	
m039.pep	130 PPATAATPAAAPQVT             PPATAATPAAAPQVT 120 130	  VPPAAPARQ	 DGFNWTIATL	 FALIVLIMQ	1111111	

Computer analysis of this amino acid sequence gave the following results: Homology with a predicted ORF from N. gonorrhoeae

ORF 039 shows 60.8% identity over a 171 aa overlap with a predicted ORF (ORF 039.ng) from N. gonorrhoeae:

m039/g039

m039.pep	10 MPSEPPYASDGIKPD	1:	:     :	111:111		
g039	MPSEPPAASDGIKPT	HTEKTSCPPVS	/RTAKPASGS	KKPSSTSPI	YASSSAWMAW	ממז צחצ
	10	20	30	40	50	60
m039.pep	70 XXXXXXXXXXXXXXXXXX	80	90	100	110	120
	XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX	:	LIVHRIGIRA	IGKKOISRI	DEIAGILNGG	TTQPDI
g039	TIWQARKNLYSTIG	PKLFRDV	:       LVHRIGTHA	:   :    .ISKKQMSRI	 EIADILNGG'	 TTT.HDT
	70	80	90	100		
<b></b>	130	140	150	160	170	
m039.pep	PPATAAT-PAAAPQVT		FNWTIATLF	ALIVLIMQL	SYLVILX	
g039	PPATAAAAPAAAPOVS		LNWTIATIF	1	CVIETTY	
	120 130	14		50	160	

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 147>:

1	ATGAACGCGC	CCGACAGCTT	TGTCGCCCAC	MMCCCCCC T T C	
51	CATCCGCCAA	ATCCCCCCA	CCACACACAC	TTCCGCGAAG	
101	TCCTCCAACC	AIGCGCGGCA	CGACACTGGT	CGCCGGCATA	GAcggCCGCC
	TGCTCGAAGG	CGGCACCTTA	AATAAGCTCG	CCGCCGACAT	CGGGCTGTTG
151	TCGCAACTGG	GCATCCGACT	CGTCCTCATC	CACGGCGCGT	ACCACTTCCT
201	CGAccgCCTC	GCCGCCGCGC	AAGgccGCAC	GCCGCATTAT	TGCCGgggtt
251	TGCGCGTTAC	CGACGaAACc	tcGctcgGAC	AGGCGCAGCA	
301	AccgTCCGCA	GCCGTTTTGA	agcCGCATTG		
351	cacaCGCGCG	CCTTCCGTCC	CCCTCCTA+c	ageggeages	tttcaggatt ctgacCGCCC
401	GTCcgatggg	Catasttasc	CGCICGIALC	gggcaacttc	ctgacCGCCC
451	CCCEERRORS	ACACCCCCC	ggaACCGata	tggaatacgc	gggggttatc
501	CCMCMCCA	ACACCGCCGC	CCTCCGTTTC		CGGGCAATAT
	CGTCTGGATG	CCGCCGCTCG	GGCATTCCTA	CCCCCCAAA	ACCTTCAATC
551	TCGATATGGT	GCAGGCCGCC	GCTTCCGTCG	CCGTCTCGCT	TCACCCCCAA
601	AAACTCGTTT	ACCTGACCCT	TTCAGACGGC	ATTTCCCGCC	CCGACGGCAC
651	GCTCGCCGAA	ACCCTCTCGG	CACAGGAAGC	GCAATCGCTG	
701	CCGCCAGCGA	AACCCGACGA	СТСАТТТССТ		GCGGAACACG
751	GGCGGCGTGC	ATCGCGTCCA	NATICOTO NA C	CCGCCGTTGC	CGCGCTCGAA
801	GCTGCAAGAA	CTCTTCACCC	COLLOCATO	GGGGCCGCCG	ACGGCAGCCT
851	y y c c c mm c c m	CTCTTCACCC	GCAACGGCAT	CGGCACGTCC	
901	AAGCCTTCGT	CTCCATCCGG	CAGGCGCACA	GCGGCGACAT	CCCGCACATC
	GCCGCCCTCA	TCCGCCCGCT	GGAAGAACAG	GGCGTCCTAT	TGCACCGCAG
951	CCGCGAATAC	CTCGAAAACC	ACATTTCCGA	ATTTTCCATC	CTCGAACACG
					O1 CGMACACG

```
1001 ACGGCGACCT GTACGGCTGT GCCGCACTCA AAACCTTTGC CGAAGCCGAT
1051 TGCGGCGAAA TCGCCTGCCT TGCCGTCTCG CCGCAGGCAC AGGACGGCGG
1101 CtACGGCGAA CGCCTGCTTG CCCACATTAT CGATAAGGCG CGCGGCATAG
1151 GCATAAGCAG GCTGTTCGCA CTGTCCACAA ATACCGGCGA ATGGTTTGCC
1201 GAACGCGGCT TTCAGACGGC ATCGGAAGAC GAGCTGCCCG AAACGCGGCG
1251 CAAAGACTAC CGCAGCAACG GACGAAACCC GCATATTCTG GTGCGTCGCC
1301 TGCACCGCTG A
```

### This corresponds to the amino acid sequence <SEQ ID 148; ORF 040.ng>: g040.pep

```
1 MNAPDSFVAH FREAAPYIRQ MRGTTLVAGI DGRLLEGGTL NKLAADIGLL
51 SQLGIRLVLI HGAYHFLDRL AAAQGRTPHY CRGLRVTDET SLGQAQQFAG
101 TVRSRFEAAL CGSVSGFARA PSVPLVSGNF LTARPMGVID GTDMEYAGVI
151 RKTDTAALRF QLDAGNIVWM PPLGHSYGGK TFNLDMVQAA ASVAVSLQAE
201 KLVYLTLSDG ISRPDGTLAE TLSAQEAQSL AEHAASETRR LISSAVAALE
251 GGVHRVQILN GAADGSLLQE LFTRNGIGTS IAKEAFVSIR QAHSGDIPHI
301 AALIRPLEEQ GVLLHRSREY LENHISEFSI LEHDGDLYGC AALKTFAEAD
351 CGEIACLAVS PQAQDGGYGE RLLAHIIDKA RGIGISRLFA LSTNTGEWFA
401 ERGFQTASED ELPETRKDY RSNGRNPHIL VRRLHR*
```

### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 149>: m040.seq

```
ATGAGCGCGC CCGACCTCTT TGTCGCCCAC TTCCGCGAAG CCGTCCCCTA
   1
  51
     CATCCGCCAA ATGCGCGGCA AAACGCTGGT CGCCGGCATA GACGACCGCC
      TGCTCGAAGG TGATACCTTA AACAAGCTCG CCGCCGACAT CGGGCTGTTG
 101
 151
     TCGCAACTGG GCATCAGGCT CGTCCTCATC CACGGCGCGC GCCACTTCCT
     CGACCGCCAC GCCGCCGCTC AAGGCCGCAC GCCGCATTAT TGCCGGGGCT
 201
      TGCGCGTTAC CGACGAAACC TCGCTCGAAC AGGCGCAGCA GTTTGCCGGC
 301
     ACCGTCCGCA GCCGTTTTGA AGCCGCATTG TGCGGCAGCG TTTCCGGGTT
 351 CGCGCGCGC CCTTCCGTCC CGCTCGTATC GGGCAACTTC CTGACCGCCC
 401 GTCCGATAGG TGTGATTGAC GGAACCGATA TGGAATACGC GGGCGTTATC
     CGCAAAACCG ACACCGCCGC CCTCCGTTTC CAACTCGACG CGGGCAATAT
 451
     CGTCTGGCTG CCGCCGCTCG GACATTCCTA CAGCGGCAAG ACCTTCTATC
 501
     TCGATATGCT TCAAACCGCC GCCTCCGCCG CCGTCTCGCT TCAGGCCGAA
 551
 601 AAACTCGTTT ACCTGACCCT TTCAGACGGC ATTTCCCGCC CCGACGGCAC
     GCTCGCCGAA ACCCTCTCGG CACAGGAAGC GCAATCGCTG GCGGAACACG
 701 CCGGCGGGCA AACGCGACGG CTGATTTCGT CCGCCGAACT CTTCACCCGC
 751 AACGGCATCG GCACGTCCAT TGCCAAAGAA GCCTTCGTCT CCATCCGGCA
     rGCGCAywgG G.CGACATCC CGCACATCGC CGCCCTCATC CGCCCGCTGG
 801
 851 AAGAACAGGG CATCCTGCTG CACCGCAs.c GCGAATACCT CGAAAACCAC
 901 ATTTCCGAAT TTTCCATCCT CGAACACGAC GGCAACCTGT ACGGTTGCGC
 951 CGCCCTGAAA ACCTTTGCCG AAGCCGATTG CGGCGAAATC GCCTGCCTTG
1001
     CCGTCTCGCC GCag.cACAG GACGGCGGCT ACGGCGAACG CnTGCTTGCC
     CACATTATCG ATAAGGCGCG CGGCATAGGC ATAAGCAGGC TGTTCGCACT
1101
     GTCCACAAAT ACCGGCGAAT GGTTTGCCGA ACGCGGCTTT CAGACGGCAT
1151
     CGGAAGACGA GTTGCCCGAA ACGCGGCGCA AAGACTACCG CAGCAACGGA
     CGGAACTCGC ATATTCTGGT ACGTCGCCTG CACCGCTGA
```

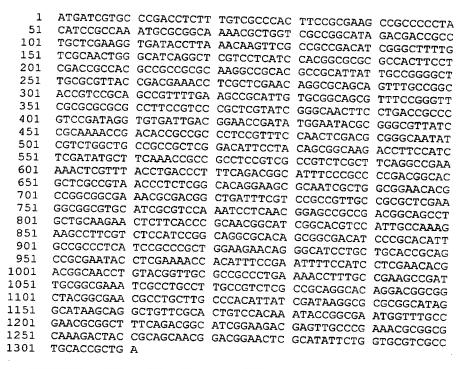
#### This corresponds to the amino acid sequence <SEQ ID 150; ORF 040>:

```
m040.pep

1 MSAPDLFVAH FREAVPYIRQ MRGKTLVAGI DDRLLEGDTL NKLAADIGLL
51 SQLGIRLVLI HGARHFLDRH AAAQGRTPHY CRGLRVTDET SLEQAQQFAG
101 TVRSRFEAAL CGSVSGFARA PSVPLVSGNF LTARPIGVID GTDMEYAGVI
151 RKTDTAALRF QLDAGNIVWL PPLGHSYSGK TFYLDMLQTA ASAAVSLQAE
201 KLVYLTLSDG ISRPDGTLAE TLSAQEAQSL AEHAGGQTRR LISSAELFTR
251 NGIGTSIAKE AFVSIRQAHX XDIPHIAALI RPLEEQGILL HRXREYLENH
301 ISEFSILEHD GNLYGCAALK TFAEADCGEI ACLAVSPQXQ DGGYGERXLA
351 HIIDKARGIG ISRLFALSTN TGEWFAERGF QTASEDELPE TRRKDYRSNG
401 RNSHILVRRL HR*
```

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 151>:

BNSDOCID: <WO___9957280A2_I_>



# This corresponds to the amino acid sequence <SEQ ID 152; ORF 040.a>:

1 MIVPDLFVAH FREAAPYIRQ MRGKTLVAGI DDRLLEGDTL NKFAADIGLL
51 SQLGIRLVLI HGARHFLDRH AAAQGRTPHY CRGLRVTDET SLEQAQQFAG
101 TVRSRFEAAL CGSVSGFARA PSVPLVSGNF LTARPIGVID GTDMEYAGVI
151 RKTDTAALRF QLDAGNIVWL PPLGHSYSGK TFHLDMLQTA ASVAVSLQAE
201 KLVYLTLSDG ISRPDGTLAV TLSAQEAQSL AEHAGGETRR LISSAVAALE
251 GGVHRVQILN GAADGSLLQE LFTRNGIGTS IAKEAFVSIR QAHSGDIPHI
301 AALIRPLEEQ GILLHRSREY LENHISEFSI LEHDGNLYGC AALKTFAEAD
351 CGEIACLAVS PQAQDGGYGE RLLAHIIDKA RGIGISRLFA LSTNTGEWFA
401 ERGFQTASED ELPETRKDY RSNGRNSHIL VRRLHR*

#### m040/a040 91.5% identity in 436 aa overlap

m040.pep	10 20 30 40 50 60 MSAPDLFVAHFREAVPYIRQMRGKTLVAGIDDRLLEGDTLNKLAADIGLLSQLGIRLVLI   :
m040.pep	70 80 90 100 110 120  HGARHFLDRHAAAQGRTPHYCRGLRVTDETSLEQAQQFAGTVRSRFEAALCGSVSGFARA
m040.pep	130 140 150 160 170 180 PSVPLVSGNFLTARPIGVIDGTDMEYAGVIRKTDTAALRFQLDAGNIVWLPPLGHSYSGK
m040.pep	190 200 210 220 230 240 TFYLDMLQTAASAAVSLQAEKLVYLTLSDGISRPDGTLAETLSAQEAQSLAEHAGGQTRR   :

	190	200	210	220	230 2	240
m040.pep	LISSA		111111	1111111111	· · · · · · · · · · · · · · · · · · ·	
a040	LISSAVAALEGGVHF 250	RVQILNGAADO 260	SSLLQELFTRNG 270	GIGTSIAKEAF 280		PHI 300
m040.pep	280 290 AALIRPLEEQGILLF		310 SEFSILEHDGN	320 LYGCAALKTF	330 AEADCGEIACLA	¥VS
a040	AALIRPLEEQGILLE 310	RSREYLENHI 320	SEFSILEHDGN 330	LYGCAALKTF 340		VS 360
m040.pep	340 350 PQXQDGGYGERXLAH	11111111	11111111111	111111111	11111111111	Ш
2040	PQAQDGGYGERLLAH 370	380	390	400		(DY 120
m040.pep	400 410 RSNGRNSHILVRRLH             RSNGRNSHILVRRLH	11				
	430					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae
ORF 040 shows 88.3% identity over a 436 aa overlap with a predicted ORF (ORF 040.ng) from N. gonorrhoeae:

#### m040/g040

m040.pep	MSAPDLFVAHFREAVPYIRQMRGKTLVAGIDDRLLEGDTLNKLAADIGLLSQLGIRLVLI	60
g040	MNAPDSFVAHFREAAPYIRQMRGTTLVAGIDGRLLEGGTLNKLAADIGLLSQLGIRLVLI	60
m040.pep	HGARHFLDRHAAAQGRTPHYCRGLRVTDETSLEQAQQFAGTVRSRFEAALCGSVSGFARA	120 "
g <b>04</b> 0	HGAYHFLDRLAAAQGRTPHYCRGLRVTDETSLGQAQQFAGTVRSRFEAALCGSVSGFARA	120
m040.pep	PSVPLVSGNFLTARPIGVIDGTDMEYAGVIRKTDTAALRFQLDAGNIVWLPPLGHSYSGK	180
g040	PSVPLVSGNFLTARPMGVIDGTDMEYAGVIRKTDTAALRFQLDAGNIVWMPPLGHSYGGK	180
m040.pep	TFYLDMLQTAASAAVSLQAEKLVYLTLSDGISRPDGTLAETLSAQEAQSLAEHAGGQTRR	240
g040	TFNLDMVQAAASVAVSLQAEKLVYLTLSDGISRPDGTLAETLSAQEAQSLAEHAASETRR	240
m040.pep	LISSAELFTRNGIGTSIAKEAFVSIRQAHXXDIPHI	276
g040		300
m040.pep	AALIRPLEEQGILLHRXREYLENHISEFSILEHDGNLYGCAALKTFAEADCGEIACLAVS	336
g040		360
m040.pep	POXODGGYGERXLAHIIDKARGIGISRLFALSTNTGEWFAERGFQTASEDELPETRRKDY	396
g040		420
m040.pep	RSNGRNSHILVRRLHRX 413	
g040		

BNSDOCID: <WO___9957280A2_l_>

LLL.PARTER_LL + PROCERTE

```
The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 153>:
```

```
1 ATGAGTTCGC CCAAACACAT CGGCTTGCAG GGCGGCAGCA ACGGCGGCCT
51 GATTACCGCC GCCGCCTTCG TGCGCGAACC GCAAAGCATC GGTGCGCTGG
101 TGTGCGAAGT ACCGCTGACC GATATGATCC GTTATCCGCT GCTGTCCGCC
151 GGTTCAAGTT GGACGGACGA ATACGGCAAT CCGCAGAAAT ACGAAGCCTG
201 CAAACGCCGG CTGGCCAAT TGTCGCCGTA TCACAATCTT TCAGACGGCA
251 TCGATTATCC GCCCGCACTC ATTACCACCA GCCTCAGCGA CGACCGCGTC
301 CATCCCGCCC ACGCGCTCAA ATTCTACGCC AAACTGCGCG AAACCTCGCC
351 GCAATCTTGG CTCTACTCGC CTGACGGCGG CGGCCATACC GGCAACGGCA
401 CCCAACGCGA ATCCGCCGAC AAACTCGCCT GCGTGTTGCT GTTTTTGAAA
```

## This corresponds to the amino acid sequence <SEQ ID 154; ORF 041.ng>:

1 MSSPKHIGLO GGSNGGLITA AAFVREPOSI GALVCEVPLT DMIRYPLLSA
51 GSSWTDEYGN POKYEACKRR LGELSPYHNL SDGIDYPPAL ITTSLSDDRV
101 HPAHALKFYA KLRETSPOSW LYSPDGGGHT GNGTORESAD KLACVLLFLK

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 155>:

1 ATCAGTTCGC CCGAACACAT CGGCTTGCAG GGCGGCAGCA ACGGCGGACT
51 GATTACTGCC GCCGCCTTCG TGCGCGAACC GCAAAGCATC GGCGCGCTGG
101 TGTGCGAAGT GCCGCTGACC GACATGATCC GTTATCCGCT GCTCTCCGCC
151 GGTTCAAGCT GGACAGACGA ATACGGCAAT CCGCAAAAAT ACGAAGTCTG
201 CAAACGCCGG TTGGGCGAAT TGTCGCCGTA TCACAATCTT TCAGACGGCA
251 TCGATTATCC GCCCGCGCTC ATTACCACCA GCCTGTCCGA CGATCGCGTC
301 CATCCCGCCC ACGCGCTCAA GTTCTACGCC AAACTGCGCG AAACCTCCGC
351 GCAATCTTGG CTCTACTCGC CTGACGGCGG CGGCCATACC GCCAACGGCA
401 CCCAACGCGA ATCCGCCGAC GAACTCGCCT GCGTCTTGCT GTTTTTGAAA

### This corresponds to the amino acid sequence <SEQ ID 156; ORF 041>:

m041.pep

1 ISSPEHIGLO GGSNGGLITA AAFVREPQSI GALVCEVPLT DMIRYPLLSA
51 GSSWTDEYGN PQKYEVCKRR LGELSPYHNL SDGIDYPPAL ITTSLSDDRV
101 HPAHALKFYA KLRETSAQSW LYSPDGGGHT GNGTQRESAD ELACVLLFLK
151 EFLG*

### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 157>:

a041.seq

1 ATCAGTTCGC CCGAACACAT CGGCTTGCAG GGCGGCAGCA ACGGCGGACT
51 GATTACTGCC GCCGCCTTCG TGCGCGAACC GCAAAGCATA GGCGCGCTGG
101 TGTGCGAAGT GCCGCTGACC GACATGATCC GTTATCCGCT GCTCTCCGCC
151 GGTTCAAGCT GGACAGACGA ATACGGCAAT CCGCAAAAAT ACGAAGTCTG
201 CAAACGCCGG TTGGGCGAAT TGTCGCCGTA TCACAATCTT TCAGACGGCA
251 TCGATTATCC GCCCGCGCTC ATTACCACCA GCCTGTCCGA CGATCGCGTC
301 CATCCCGCCC ACGCGCTCAA GTTCTACGCC AAACTGCGCG AAACTCGCCC
351 GCAATCTTGG CTCTACTCGC CTGACGGCGG CGGCCATACC GGCAACGGCA
401 CGCAGCGCGA AGCCGCCGAC GAACTCGCCT GCGTGTTGCT GTTTTTGAAA
451 GAGTTTTTGG GCTAA

## This corresponds to the amino acid sequence <SEQ ID 158; ORF 041.a>:

1 ISSPEHIGLQ GGSNGGLITA AAFVREPQSI GALVCEVPLT DMIRYPLLSA
51 GSSWTDEYGN PQKYEVCKRR LGELSPYHNL SDGIDYPPAL ITTSLSDDRV
101 HPAHALKFYA KLRETSPQSW LYSPDGGGHT GNGTQREAAD ELACVLLFLK

151 EFLG*

#### m041/a041 98.7% identity over a 154 aa overlap

	10	20	30	40	50	60
m041.pep	ISSPEHIGLQGGSN	GGLITAAAF	VREPQSIGALV	CEVPLTDMIF	RYPLLSAGSSW	TDEYGN
	11111111111111			11111111111	11111111111	111111
a041	ISSPEHIGLQGGSN	GGLITAAAF	VREPQSIGALV	CEVPLTDMIR	YPLLSAGSSW	TDEYGN
	10	20	30	40	50	60
	70	80	90	100	110	120
m041.pep	PQKYEVCKRRLGEL	SPYHNLSDG:	IDYPPALITTS:	LSDDRVHPAH	IALKFYAKLRE	TSAQSW
				1111111111	111111111	11 111
a041	PQKYEVCKRRLGEL.	SPYHNLSDG	IDYPPALITTS	LSDDRVHPAH	ALKFYAKLRE	TSPQSW
	70	80	90	100	110	120
	130	140	150			
m041.pep	LYSPDGGGHTGNGT	QRESADELA	CVLLFLKEFLG	X		
	111111111111	111:1111		I		
a041	LYSPDGGGHTGNGT	QREAADELAG	CVLLFLKEFLG	Κ.		
	130	140	150			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae

ORF 041 shows 96.8% identity over a 154 aa overlap with a predicted ORF (ORF 041.ng) from N. gonorrhoeae:

m041/g041

	10	20	30	40	50	60
m041.pep	ISSPEHIGLQGGSN	GGLITAAAF	VREPQSIGALV	CEVPLTDMIF	RYPLLSAGSSV	TDEYGN
	:       :	11111111	1111111111	1111111111	111111111111	
g041	MSSPKHIGLQGGSN	GGLITAAAF	VREPQSIGALV	CEVPLTDMIF	RYPLLSAGSSV	TDEYGN
	10	20	30	.40	50	60
	70	80	90	100	110	120
m041.pep	PQKYEVCKRRLGEL	SPYHNLSDG:	IDYPPALITTS	LSDDRVHPAH	IALKFYAKLRE	
		THIFT		1111111111	111111111	11 111
g041	PQKYEACKRRLGEL	SPYHNLSDG:	IDYPPALITTS	LSDDRVHPAH	IALKFYAKLRE	TSPOSW
	70	80	90	100	110	120
	130	140	150			
m041.pep	LYSPDGGGHTGNGT	QRESADELA(	CVLLFLKEFLG:	х		
		111111:11	11111111111	1		
g041	LYSPDGGGHTGNGT	QRESADKLA	CVLLFLKEFLG	X		
	130	140	150			

# The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 159>: g041-1.seq

т –	1.36	4				
	1	ATGAAATCCT	ACCCCGACCC	CTACCGCCAT	TTTGAAAACC	TCGATTCCGC
	51	CGAAACGCAA	AACTTCGCTG	CTGAAGCGAA	TGCCGAAACG	CGCGCGCGTT
	101		CGACAAGGCG			
	151		CGCGGCAGAT			
	201	GTACCATTTC	CATCAGAATG	CGGAATATCC	GAAGGGCGTG	TACCGCATGT
	251	GTACGGCGGC	GACCTACCGT	TCCGGCTATC	CCGAGTGGAA	AATCCTGTTT
	301	TCGGTGGCGG	ATTTCGATGA	GTTGCTCGGC	GACGATGTGT	ATTTGGGCGG
	351	CGTGTCGCAC	TTGGTGGAGC	AGCCCAACCG	CGCGCTGCTG	ACTTTGAACA
	401	AATCGGGCGG	CGATACGGCG	TATACGCTGG	AAGTGGATTT	GGAAGCAGGG
	451	GAATTGGTAG	AGGGCGGTTT	TCACTTTCCG	GCAGGCAAAA	ACCATGTGTC
	501		GAAAACAGCG			
	551	AGTTGACCGA	ATCGGGCTAT	CCGCGCGAAG	TGTGGCTGGT	GGAACGCGGC

```
601 AAGAGTTTCG AGGAAAGCCT GCCGGCGTAC CAAATCGATA AAGGCGCGAT
      GATGGTAAAC GCGTGGCGTT ACCTCGATCC GCAGGGTTCG CCGATTGATT
      TGATTGAAGC GTCGGACGGT TTTTACACCA AGACGTATTT GCAGGTGTCG
 751 TCCGAAGGCG GGGCGAAACC GTTGAACCTG CCTAATGATT GCGATGTGGT
 801 CGGCTATCTG GCGGGACATC TTTTGCTGAC GCTGCGCAAG GACTGGCACC
     GCGCGAACCA AAGCTATCCG AGTGGCGCGT TGGTGGCGGT GAAACTGAAT
      CGGGGCGAAC TCGGGGCGGC GCAGCTTTTG TTTGCGCCCG ATGAAACGCA
     GGCATTGGAA AGCGTGGAAA CGACCAAGCG TTTTGTGGTG GCAAGCCTGC
 951
     TGGAGAATGT ACAAGGCCGT CTGAAAGCGT GGCGGTTTGC CGACAGCAAA
1051 TGGCAGGAAG CCGAGTTGCC GCACCTGCCC TCGGGCGCGT TGGAAATGAC
1101 CGACCAACCG TGGGGCGGCG ACGTGGTTTA TCTTGCCGCC AGCGATTTCA
      CCACGCCGCT GACGCTGTTT GCGCTGGATT TGAACGTGAT GGAACTGACC
1151
1201 GTCATGCGCC TCCAGCCGCA GCAGTTTGTT TCAGACGGCA TCGAAGTGCG
1251 GCAGTTTTGG GCGGTGTCGT CCGACGGCGA ACGCATTCCT TATTTCCACG
     TCGGCAAAAA CGCCGCGCCC GACACGCCGA CCTTAGTCTA TGCTTACGGA
1301
1351 GGTTTCGGCA TTCCTGAATT GCCGCATTAT CTGGGCAGCG TCGGCAAATA
1401
      TTGGCTGGAA GAGGGCAATG CCTTTGTATT GGCAAACATC CGCGGCGGCG
     GAGAATTCGG CCCGCGCTGG CATCAGGCGG CGCAGGGAAT CAGCAAACAC
1451
1501 AAAAGCGTTG ATGATTTGTT GGCAGTCGTG CGTGATTTGT CCGAACGCGG
1551 CATGAGTTCG CCCAAACACA TCGGCTTGCA GGGCGGCAGC AACGGCGGCC
1601 TGATTACCGC CGCCGCCTTC GTGCGCGAAC CGCAAAGCAT CGGTGCGCTG
1651 GTGTGCGAAG TACCGCTGAC CGATATGATC CGTTATCCGC TGCTGTCCGC
1701
     CGGTTCAAGT TGGACGGACG AATACGGCAA TCCGCAGAAA TACGAAGCCT
1751 GCAAACGCCG GCTGGGCGAA TTGTCGCCGT ATCACAATCT TTCAGACGGC
1801 ATCGATTATC CGCCCGCACT CATTACCACC AGCCTCAGCG ACGACCGCGT
1851 CCATCCCGCC CACGCGCTCA AATTCTACGC CAAACTGCGC GAAACCTCGC
     CGCAATCTTG GCTCTACTCG CCTGACGGCG GCGGCCATAC CGGCAACGGC
     ACCCAACGCG AATCCGCCGA CAAACTCGCC TGCGTGTTGC TGTTTTTGAA
1951
2001 AGAATTTTTG GGATAA
```

# This corresponds to the amino acid sequence <SEQ ID 160; ORF 041-1.ng>:

```
MKSYPDPYRH FENLDSAETQ NFAAEANAET RARFLNNDKA RALSDGILNQ
  1
 51 MQDTRQIPFC QEHRARMYHF HQNAEYPKGV YRMCTAATYR SGYPEWKILF
    SVADFDELLG DDVYLGGVSH LVEQPNRALL TLNKSGGDTA YTLEVDLEAG
151 ELVEGGFHFP AGKNHVSWRD ENSVWVCPAW DERQLTESGY PREVWLVERG
201 KSFEESLPAY QIDKGAMMVN AWRYLDPQGS PIDLIEASDG FYTKTYLQVS
     SEGGAKPLNL PNDCDVVGYL AGHLLLTLRK DWHRANQSYP SGALVAVKLN
    RGELGAAQLL FAPDETQALE SVETTKRFVV ASLLENVQGR LKAWRFADSK
301
    WQEAELPHLP SGALEMTDQP WGGDVVYLAA SDFTTPLTLF ALDLNVMELT
351
    VMRLQPQQFV SDGIEVRQFW AVSSDGERIP YFHVGKNAAP DTPTLVYAYG
401
451 GFGIPELPHY LGSVGKYWLE EGNAFVLANI RGGGEFGPRW HQAAQGISKH
    KSVDDLLAVV RDLSERGMSS PKHIGLQGGS NGGLITAAAF VREPQSIGAL
551
    VCEVPLTDMI RYPLLSAGSS WTDEYGNPQK YEACKRRLGE LSPYHNLSDG
    IDYPPALITT SLSDDRVHPA HALKFYAKLR ETSPQSWLYS PDGGGHTGNG
    TORESADKLA CVLLFLKEFL G*
651
```

## The following partial DNA sequence was identified in N. meningitidis <SEQ ID 161>:

1.26	: ·				
1	ATGAAATCCT	ACCCCGACCC	CTACCGCCAT	TTTGAAAACC	TCGATTCCGC
51	CGAAACGCAA	AACTTCGCTG	CTGAAGCGAA		CGCGCGCGTT
101	TTTTAGAAAA	CGACAAGGCG	CGCGCGCTTT		TTTGGCGCAG
151	TTGCAGGACA	CGCGGCAGAT	TCCGTTTTGT	CAGGAACACC	GCGCGCGCAT
201	GTACCATTTC	CATCAGGACG	CGGAGTATCC	GAAGGGCGTG	TACCGCGTGT
251	GTACCGCGGC	GACGTATCGT	TCCGGCTATC	CCGAGTGGAA	AATCCTGTTT
301	TCGGTGGCGG	ATTTCGACGA	ATTGCTTGGC	GACGATGTGT	ATTTGGGCGG
351	CGTGTCGCAC	TTGGTGGAAC	AGCCCAACCG		ACACTGAGCA
401	AATTGGGCAG	CGATACGGCG	TACACGCTGG		GGAAGCAGGG
451	GAGTTGGTCG	AAGGCGGTTT	TCACTTTCCG		ACCATGTGTC
501	GTGGCGCGAT	GAAAACAGCG	TGTGGGTGTG		AACGAACGCC
551	AGTTGACCCA	ATCGGGCTAT	CCGCGCGAAG		GGAACGCGGC
601		AGGAAAGCCT		CAAATCGGCG	AAGACGCCAT
651	GATGGTGAAC	GCGTGGCGTT	ATCTCGATCC	GCAGGGTTCG	CCGATTGATT
701	TGATTGAAGC	GTCGGACGGT	TTTTACACCA		GCGGGTCTCA
					CCCCCTCTCA

1 1 3 3 1 1 2 2 th 1 777 ats

751	GCCGAAGGCG	AGGCGAAACC	GTTAAACCTG	CCCAACGATT	GCGACGTGGT
801	CGGCTATCTG	GCGGGGCATC	TTTTGCTGAC	GCTGCGCAAG	GACTGGAACC
851	GCGCGAACCA	AAGCTATCCG	AGCGGCGCGC	TGGTGGCGGT	GAAGCTGAAT
901	CGGGGCGAAC	TCGGGGCGGC	GCAGCTTTTG	TTTGCGCCCG	ATGAAACGCA
951	GGCATTGGAA	AGCGTGGAAA	CGACCAAGCG	TTTTGTGGTG	GCGAGCCTGT
1001	TGGAGAACGT	ACAAGGCCGT	CTGAAAGCAT	GGCGGTTTGC	CGACGGCAAA
1051	TGGCAGGAAG	TCGAATTGCC	GCGCCTGCCT	TCGGGCGCGT	TGGAAATGAC
1101	CGACCAACCT	TGGGGCGGCG	ACGTGGTTTA	CCTTGCCGCC	AGCGATTTCA
1151	CCACGCCGCT	GACGCTGTTT	GCGCTGGATT	TGAACGTGAT	GGAACTGACC
1201	GTCATGCGCC	GCCAGCCGCA	GCAGTTTGAT	TCAGACGGCA	TTAACGTGCA
1251	GCAGTTTTGG	ACGACTTCGG	CTGACGGCGA	GCGCATTCCT	TATTTCCACG
1301	TCGGCAAAAA	CGCCGCGCCC	GACATGCCGA	CGCTGGTCTA	TGCCTACGGC
1351	GGTTTCGGCA	TTCCCGAATT	GCCGCATTAT	CTGGGCAGCA	TTGGCAAATA
1401	TTGGCTGGAA	GAGGGCAATG	CCTTTGTATT	GGCGAACATC	CGCGGCGGCG
1451	GCGAGTTCGG	CCCGCGCTGG	CATCAGGCGG	CGCAGGGAAT	CAGCAAACAT
1501	AAAAGCGTTG	ATGATTTATT	GGCAGTCGTG	CGCGATTTGT	CCGAACGCGG
1551	TATCAGTTCG	CCCGAACACA	TCGGCTTGCA	GGGCGGCAGC	AACGGCGGAC
1601	TGATTACTGC	CGCCGCCTTC	GTGCGCGAAC	CGCAAAGCAT	CGGCGCGCTG
1651	GTGTGCGAAG	TGCCGCTGAC	CGACATGATC	CGTTATCCGC	TGCTCTCCGC
1701	CGGTTCAAGC	TGGACAGACG	AATACGGCAA	TCCGCAAAAA	TACGAAGTCT
1751	GCAAACGCCG	GTTGGGCGAA	TTGTCGCCGT	ATCACAATCT	TTCAGACGGC
1801	ATCGATTATC	CGCCCGCGCT	CATTACCACC	AGCCTGTCCG	ACGATCGCGT
1851	CCATCCCGCC	CACGCGCTCA	AGTTCTACGC	CAAACTGCGC	GAAACCTCCG
1901	CGCAATCTTG	GCTCTACTCG	CCTGACGGCG	GCGGCCATAC	CGGCAACGGC
1951	ACCCAACGCG	AATCCGCCGA	CGAACTCGCC	TGCGTCTTGC	TGTTTTTGAA
2001	AGAGTTTTTG	GGCTAA			

# This corresponds to the amino acid sequence <SEQ ID 162; ORF 041-1>: m041-1.pep

	7					
1	MKSYPDPYRH	FENLDSAETQ	NFAAEANAET	RARFLENDKA	RALSDGILAQ	
51	LQDTRQIPFC	QEHRARMYHF	HQDAEYPKGV	YRVCTAATYR	SGYPEWKILF	
101	SVADFDELLG	DDVYLGGVSH	LVEQPNRALL	TLSKLGSDTA	YTLEVDLEAG	
151	ELVEGGFHFP	AGKNHVSWRD	ENSVWVCPAW	NERQLTQSGY	PREVWLVERG	
201	KSFEESLPVY	QIGEDGMMVN	AWRYLDPQGS	PIDLIEASDG	FYTKTYLRVS	
251	AEGEAKPLNL	PNDCDVVGYL	AGHLLLTLRK	DWNRANQSYP	SGALVAVKLN	
301	RGELGAAQLL	FAPDETQALE	SVETTKRFVV	ASLLENVQGR	LKAWRFADGK	
351	WQEVELPRLP	SGALEMTDQP	WGGDVVYLAA	SDFTTPLTLF	ALDLNVMELT	
401	VMRRQPQQFD	SDGINVQQFW	TTSADGERIP	YFHVGKNAAP	DMPTLVYAYG	
451	GFGIPELPHY	LGSIGKYWLE	EGNAFVLANI	RGGGEFGPRW	HQAAQGISKH	
501	KSVDDLLAVV	RDLSERGISS	PEHIGLQGGS	NGGLITAAAF	VREPQSIGAL	
551	VCEVPLTDMI	RYPLLSAGSS	WTDEYGNPQK	YEVCKRRLGE	LSPYHNLSDG	
601	IDYPPALITT	SLSDDRVHPA	HALKFYAKLR	ETSAQSWLYS	PDGGGHTGNG	
651	TORESADELA	CVLLFLKEFL	G*			

#### m041-1/g041-1 94.6% identity in 671 aa overlap

	10	20	30	40	50	60
m041-1.pep	MKSYPDPYRHFENLI	SAETQNFAA:	EANAETRARF	LENDKARALS	DGILAQLQDT	RQIPFC
			1111111111	1:1111111	1111 1:111	11111
g041-1	MKSYPDPYRHFENLI	DSAETQNFAA:	EANAETRARF	LNNDKARALS	DGILNQMQDI	RQIPFC
	10	20	30	40	50	60
	70	80	90	100	110	120
m041-1.pep	QEHRARMYHFHQDAI	EYPKGVYRVC	TAATYRSGYF	EWKILFSVAD	FDELLGDDVY	LGGVSH
			111111111	1111111111	1111111111	11111
g041-1	QEHRARMYHFHQNAI	EYPKGVYRMC	TAATYRSGYF	EWKILFSVAD	FDELLGDDVY	LGGVSH
	70	80	90	100	110	120
	130	140	150	160	170	180
m041-1.pep	LVEQPNRALLTLSKI	LGSDTAYTLE	VDLEAGELVE	GGFHFPAGKN	HVSWRDENSV	WVCPAW
		1:111111	111111111	111111111	HILLIAM	11111
g041-1	LVEQPNRALLTLNKS			GGFHFPAGKN	HVSWRDENSV	WVCPAW
	130	140	150	160	170	180
	190	200	210	220	230	240
m041-1.pep	NERQLTQSGYPREVV	VLVERGKSFE		DGMMVNAWRY	LDPQGSPIDI	IEASDG
	:		1111:111:	: [ ] ] [ ] [ ]	111111111	

BNSDOCID: <WO___9957280A2_J_>

190   200   210   220   230   240   240   250   260   270   280   290   30   300   300   300   300   300   300   300   300   300   300   300   300   300   300   300   300   300   300   300   300   300   300   300   300   300   300   300   300   300   350   360   360   350   360   360   350   360   360   350   360   360   350   360   360   350   360   360   350   360   360   350   360   360   350   360   360   350   360   360   350   360   360   350   360   360   350   360   360   350   360   360   350   360   360   350   360   360   350   360   360   350   360   360   350   360   360   350   360   360   350   360   360   350   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360							
190   200   210   220   230   240   240   230   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240   240	g041-1	DERQLTESGYPREV	WLVERGKSF	EESLPAYQID	KGAMMVNAWR	YLDPQGSPI	LIEASDG
M041-1.pep		190	200	210	220		240
M041-1.pep		250	260	0.70			
	m041-1.pep			270	280	290	300
		IIIIIII:II:II	HILLILL	LILLILLI	FFTLTKKDMN	RANOSYPSGA	LVAVKLN
Mod1-1.pep	g041-1	FYTKTYLOVSSEGG	AKPINI.PND	11111111111111111111111111111111111111	 		
310   320   330   340   350   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360   360		250	260	270	280		
M041-1.pep   RGELGAAQLIFAPDETQALESVETTKREVVASLLENVQCRIKAWRFADGKWQEVELPRI				2.0	200	290	300
RELIGAQLIFAPDETOALESVETTKREVVASLLENVQGRLKAWRFADGKWQEVELPRL   RGELGAAQLIFAPDETQALESVETTKREVVASLLENVQGRLKAWRFADGKWQEVELPRL   310 320 330 340 350 36				330	340	350	360
G041-1   RGELGAAQLIFAPDETQALESVETKREVVASLLENVQGRLKAWRFADSKWQEAELPHL   310   320   330   340   350   360   360   360   370   380   390   400   410   42   42   42   42   42   42   42   4	m041-1.pep	RGELGAAQLLFAPD	ETQALESVE	TTKRFVVASL	LENVQGRLKAI	MREADCRMOR	TEL DOLD
310   320   330   340   350   360   360   360   370   380   390   400   410   42   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   440   450   460   470   480   430   430   440   450   460   470   480   430   430   440   450   460   470   480   430   430   440   450   460   470   480   430   440   450   460   470   480   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430   430	0 4 2 - 1		1 1 1 1 1 1 1 1 1			1111-111	
MO41-1.pep   SGALEMTDOPWGGDVVYLAASDFTTPLTLFALDLNVMELTVMRQPQOFDSDGINVQQF   SGALEMTDOPWGGDVVYLAASDFTTPLTLFALDLNVMELTVMRQPQOFDSDGINVQQF   SGALEMTDOPWGGDVVYLAASDFTTPLTLFALDLNVMELTVMRQPQOFVSDGINVQQF   370 380 390 400 410 42	g041-1	RGELGAAQLLFAPD.	ETQALESVE'	TTKRFVVASLI	LENVQGRLKA	WRFADSKWOE	AELPHLP
M041-1.pep   SGALEMTDQPWGGDVVYLAASDFTTPLTLFALDLNVMELTVMRRQPQQFDSDGINVQQF		310	320	3.30	340		360
M041-1.pep   SGALEMTDQPWGGDVVYLAASDFTTPLTLFALDLNVMELTVMRRQPQQFDSDGINVQQF		270	200				
SGALEMTDQPWGGDVYLAASDFTTPLTLFALDLNVMELTVMRLQPQQFVSDGIEVRQFF   370 380 390 400 410 42	m041-1.pep		380	390	400	410	420
370   380   390   400   410   42    42    430   440   450   460   470   48    48    48    450   460   470   48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48    48		I I I I I I I I I I I I I I I I I I I	VVILAASDE;	TTPLTLFALDI	LNVMELTVMRF	ROPOOFDSDG	INVQQFW
### ### ##############################	g041-1	SGALEMTDOPWGGD	VVYI. A A S D E "	ן ן ן ן ן ן ן ן ן ן ן ן ן ן ן ן ן ן ן		11111111	1:1:11
### ### ##############################	•	370	380	390	VOO TAMETIANKI		
TTSADGERIPYFHVGKNAAPDMPTLVYAYGGFGIPELPHYLGSIGKYWLEEGNAFVLAN:   1				330	400	410	420
TTSADGERIPYFHVGKNAAPDMPTLVYAYGGFGIPELPHYLGSIGKYWLEEGNAFVLAN:				450	460	470	400
AVSSDERIPYFHVGKNAAPDTPTLVAYGGGFGIPELPHYLGSVGKYWLEEGNAFVLAN:   430	m041-1.pep	TTSADGERIPYFHV(	GKNAAPDMPI	LVYAYGGFGI	PELPHYLEST	CKYWI FECNI	V 124 21 20 91 2
### AVSSIGERTPYFHVGKNAAPDTPTLVYAYGGFGIPELPHYLGSVGKYWLEEGNAFVLAN:  #### 430				111111111	111111111	3 ( ) 1 1 1 1 1 1 1 1	
### ### ##############################	g041-1	AASSDGEKILLLHAG	SKNAAPDTPI	LVYAYGGFGI	PELPHYLGSV	GKYWLEEGN	AFVLANT
### Table   RGGGEFGPRWHQAAQGISKHKSVDDLLAVVRDLSERGISSPEHIGLQGGSNGGLITAAAF		430	440	450	460		480
### Table   RGGGEFGPRWHQAAQGISKHKSVDDLLAVVRDLSERGISSPEHIGLQGGSNGGLITAAAF		400	* • •				
RGGGEFGPRWHQAAQGISKHKSVDDLLAVVRDLSERGMSSPKHIGLQGGSNGGLITAAAA	m041-1 pep		500	510	520	530	540
MGGGLFGFRWHQAAQGISKHKSVDDLLAVVRDLSERGMSSPKHIGLQGGSNGGLITAAAF   490   500   510   520   530   540   540   550   560   570   580   590   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   600   6		I I I I I I I I I I I I I I I I I I I	GISKHKSVD	DLLAVVRDLS	ERGISSPEHI	GLQGGSNGGI	ITAAAF
### ### ##############################	q041-1	RGGGEFGPRWHOAAC					11111
Mod1-1.pep   S50   S60   S70   S80   S90   G00		490	500	510	ERGMSSPKHI		
M041-1.pep   VREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGNPQKYEVCKRRLGELSPYHNLSDG				510	320	530	540
### WREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGNPQKYEVCKRRLGELSPYHNLSDG				570	580	590	600
### TITLE TO THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE PROOF OF THE P	m041-1.pep	VREPQSIGALVCEVP	LTDMIRYPL	LSAGSSWTDE	YGNPOKYEVC	KERT CET COV	LINIT CDC
The state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the state of the		_		11111111	1111111111		
m041-1.pep   G10   G20   G30   G40   G50   G60	g041-1	VKEPOSIGALVCEVP	TIDMIKABL	LSAGSSWTDE:	YGNPQKYEAC	KRRLGELSPY	HNLSDG
m041-1.pep IDYPPALITTSLSDDRVHPAHALKFYAKLRETSAQSWLYSPDGGGHTGNGTQRESADELA  g041-1 IDYPPALITTSLSDDRVHPAHALKFYAKLRETSAQSWLYSPDGGGHTGNGTQRESADKLA 610 620 630 640 650 660  670  m041-1.pep CVLLFLKEFLGX		550	560	570	580		600
m041-1.pep IDYPPALITTSLSDDRVHPAHALKFYAKLRETSAQSWLYSPDGGGHTGNGTQRESADELA  g041-1 IDYPPALITTSLSDDRVHPAHALKFYAKLRETSAQSWLYSPDGGGHTGNGTQRESADKLA 610 620 630 640 650 660  670  m041-1.pep CVLLFLKEFLGX		610	600		_		
g041-1 IDYPPALITTSLSDDRVHPAHALKFYAKLRETSPQSWLYSPDGGGHTGNGTQRESADKLA 610 620 630 640 650 660  670  m041-1.pep CVLLFLKEFLGX	m041-1.pep			630 EVAKI DEMOR	640	650	660
### ##################################			NVDEADALA.	FIAKLKETSA(	QSWLYSPDGG(	SHTGNGTORE	SADELA
610 620 630 640 650 660 670 m041-1.pep CVLLFLKEFLGX	g041-1	IDYPPALITTSLSDD	RVHPAHAT.KI	IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII			111:11
670 m041-1.pep CVLLFLKEFLGX		610	620	630	640		
m041-1.pep CVLLFLKEFLGX					040	650	660
g041-1 CVLLFLKEFLGX 670							
g041-1 CVLLFLKEFLGX 670 1/P55577	mU41-1.pep						
670 1/P55577	~041 1						
1/P55577	GO4T-I						
		6/0					
5577 Y4NA_RHISN PROBABLE PEPTIDASE Y4NA >gi!2182536 (AF000086) Y477 (Dii							
AN A LI TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELLE TELL	5577 Y4NA_RHISN	PROBABLE PEPTIDAS	E Y4NA >gi	2182536 (AE	000086) Y4n	A [Rhizobin	ITM EX

#### m041-1/

sp | P55 SPIESSS//|Y4NA_RHISN PROBABLE PEPTIDASE Y4NA >gi|2182536 (AE000086) Y4nA [Rhizobium sp. NGR234] Length = 726

Score = 370 bits (940), Expect = e-101 Identities = 217/682 (31%), Positives = 331/682 (47%), Gaps = 22/682 (3%)

- KSYPDPYRHFENLDSAETQNFAAEANAETRARFLENDKARALSDGILAQLQDTRQIPFCQ 61 K DP + +D + + N T + ++ + L LQ T +I
- Sbjct: 42 KDASDPRAYLNEIDGDKAMTWVEAHNLSTVDKLSKDPRYSEYQADALTILQATDRIASPS 101
- Query: 62 EHRARMY-HFHQDAEYPKGVYRVCTAATYRSGYPEWKILFSVADFDELLGDDVYLGGVSH 120 R M +F QD + +G++R T +YRSG P+W+ + V
- Sbjct: 102 FARDGMIDNFWQDGTHVQGLWRRTTWESYRSGNPQWRTILDVDALSKAEGKTWVFEGGDC 161
- Query: 121 LVEQPNRALLTLSKLGSDTAYTLEVDLEAGELVEGGFHFPAGKNHVSWRDENSVWVCPAW 180 L N L+ LS G D E D+ GE V+ GF P GK V+W DEN+++V

```
Sbjct: 162 LPPTSNLCLIRLSDGGKDADVVREFDIAKGEFVKEGFVLPEGKQSVTWVDENTIYVTREW 221
Query: 181 NERQLTQSGYPREVWLVERGKSFEESLPVYQ-----IGEDGMM--VNAWRYLDPQGSPI 232
              ++T SGY
                        +V+RG+S ++++ +++
                                                 E G++ ++
Sbjct: 222 TPGEVTSSGYAYVTKVVKRGQSLDQAVEIFRGQKKDVSAERGVLRDIDGKYVMDTSYRGL 281
Query: 233 DLIEASDGFYTKTYLRVSAEGEAKPLNLPNDCDVVGYLAGHLLLTLRKDWNRANQS-YPS 291
                 FY + + L LP GY G + L+ DW A + + +
Sbjct: 282 DFFNTELAFYPNGH----PDTRKVVLPLPTTAVFSGYYKGQAIYWLKSDWTSAKGTVFHN 337
Query: 292 GALVAVKLNRGELGAAQL----LFAPDETQALESVETTKRFVVASLLENVQGRLKAWRFA 347 GA++A L A++ LF P+E Q++ TK +V S+L NV +++++ F
Sbjct: 338 GAIIAFDLKAALADPARVEPLVLFMPNEHQSVAGTTQTKNRLVLSILSNVTSEVRSFDFG 397
Query: 348 DGKWQEVELPRLPSGALEMTDQPWGGDVVYLAASDFTTPLTLFALDLNVMELTVMRRQPQ 407
                                  D +++ + F P TLF D
           GW +L + L+T
Sbjct: 398 KGGWSSFKLALPENSTLSLTSSDDESDQLFVFSEGFLEPSTLFCADAATGQVEKITSTPA 457
Query: 408 QFDSDGINVQQFWTTSADGERIPYFHVGKNAAP---DMPTLVYAYGGFGIPELPHYLGSI 464
           +FD+ G+ QQFW TS DG ++PYF V + PT++YAYGGF IP P Y
Sbjct: 458 RFDAGGLQAQQFWATSKDGTKVPYFLVARKDVKLDGTNPTILYAYGGFQIPMOPSYSAVL 517
Query: 465 GKYWLEEGNAFVLANIRGGGEFGPRWHQAAQGISKHKSVDDLLAVVRDLSERGISSPEHI 524
          GK WLE+G A+ LANIRGGGEFGP+WH A ++ + DD AV +DL + ++S H+
Sbjct: 518 GKLWLEKGGAYALANIRGGGEFGPKWHDAGLKTNRQRVYDDFQAVAQDLIAKKVTSTPHL 577
Query: 525 GLQGGSNGGLITAAAFVREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGNPQKYEVC 584
           G+ GGSNGGL+
                         ++ P
                                 A+V +VPL DM+ + +SAG+SW EYG+P
Sbjct: 578 GIMGGSNGGLLMGVQMIQRPDLWNAVVIQVPLLDMVNFTRMSAGASWQAEYGSPDD-PVE 636
Query: 585 KRRLGELSPYHNLSDGIDYPPALITTSLSDDRVHPAHALKFYAKLRETSAQSWLYSPDGG 644
             L +SPYHN+ G+ YP TS DDRV P HA K A +
Sbjct: 637 GAFLRSISPYHNVKAGVAYPEPFFETSTKDDRVGPVHARKMAALFEDMGLPFYYYENIEG 696
Query: 645 GHTGNGTQRESADELACVLLFL 666
          GH
                 +E A A +++
Sbjct: 697 GHAAAANLQEHARRYALEYIYM 718
```

### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 163>: a041-1.seq

1 ATGAAATCCT ACCCCGACCC CTACCGCCAT TTTGAAAACC TCGATTCCGC CGAAACGCAA AACTTCGCTG CTGAAGCGAA TGCCGAAACG CGCGCGCGTT 51 TTTTAAACAA CGACAAGGCA CGCGCATTGT CTGACGGCAT TTTGGCGCAG 101 TTGCAGGACA CGCGGCAAAT TCCGTTTTGT CAGGAACACC GCGCGCGGAT 201 GTACCATTTC CATCAAGATG CGGAATATCC GAAAGGCGTG TACCGCGTGT GTACCGCGGC GACTTACCGT TCGGGCTATC CTGAGTGGAA AATCCTGTTT 251 301 TCGGTGGCGG ATTTCGACGA ATTGCTCGGT GACGATGTAT ATCTAGGCGG 351 CGTGTCGCAC CTGGTGGAAC AGCCCAACCG CGCGTTGTTA ACACTGAGCA 401 AATCGGGCGG CGATACCGCG TACACGCTGG AAGTGGATTT GGAAGCAGGG 451 GAGTTGGTAG AAGGCGGTTT TCACTTTCCG GCAGGCAAAA ACCATGTGTC 501 GTGGCGCGAT GAAAACAGCG TGTGGGTGTG TCCGGCTTGG GACGAACGCC AGTTGACCGA ATCGGGCTAT CCGCGCGAGG TGTGGCTGGT GGAACGCGGC 551 601 AAGAGTTTCG AGGAAAGCCT GCCGGTGTAC CAAATTGCTG AAGACGGCAT 651 GATGGTGAAC GCGTGGCGTT ACCTCGATCC GCAGGGTTCG CCGATTGATT 701 TGATTGAAGC GTCTGACGGT TTTTACACCA AAACCTATTT GCAGGTCTCA 751 GCCGAAGGCG AAGCGAAACC GTTAAACCTG CCCAACGATT GCGACGTAGT 801 CGGCTATCTG GCCGGACATC TTTTGCTGAC CTTGCGTAAA GACTGGCACC 851 GCGCGAACCA AAGCTATCCG AGTGGCGCAT TGGTAGCAGT AAAATTAAAC 901 CGCGGCGAAT TGGGCGCGGC GCAGCTTTTG TTTGCGCCCA ATGAAACGCA 951 GGCATTGGAA AGCGTGGAAA CGACCAAGCG TTTTGTCGTG GCGAGCCTGC 1001 TGGAAAACGT ACAGGGTCGT CTGAAAGCGT GGCGTTTTAC TGATGGCAAA 1051 TGGCAGGAAA CCGAGTTGCC GCGCCTGCCT TCGGGCGCGT TGGAAATGAC 1101 CGACCAACCG TGGGGGGGCG ACGTAGTTTA CCTTGCCGCC AGCGATTTCA 1151 CCACGCCGCT GACGCTGTTT GCATTGGATT TGAACGTGAT GGAACTGACC 1201 GTCATGCGCC GCCAGCCGCA GCAGTTTGAT TCAGACGGCA TTAACGTGCA 1251 GCAGTTTTGG ACGACTTCGG CTGACGGCGA GCGCATTCCT TATTTCCACG 1301 TCGGCAAAAA CGCCGCGCCC GACATGCCGA CGCTGGTCTA TGCCTACGGC GGTTTCGGCA TTCCCGAATT GCCGCATTAT CTGGGCAGCA TTGGCAAATA 1351 1401 TTGGCTGGAA GAGGGCAATG CCTTTGTATT GGCGAACATC CGCGGCGGCG

BNSDOCID: <WO__9957280A2 L>

# This corresponds to the amino acid sequence <SEQ ID 164; ORF 041-1.a>: a041-1.pep

```
1 MKSYPDPYRH FENLDSAETQ NFAAEANAET RARFLNNDKA RALSDGILAQ
51 LQDTRQIPFC QEHRARMYHF HQDAEYPKGV YRVCTAATYR SGYPEWKILF
101 SVADFDELLG DDVYLGGVSH LVEQPNRALL TLSKSGGDTA YTLEVDLEAG
151 ELVEGGFHFP AGKNHVSWRD ENSVWVCPAW DERQLTESGY PREVWLVERG
201 KSFEESLPVY QIAEDGMMVN AWRYLDPQGS PIDLIEASDG FYTKTYLQVS
251 AEGEAKPLNL PNDCDVVGYL AGHLLLTLRK DWHRANQSYP SGALVAVKLN
     RGELGAAQLL FAPNETQALE SVETTKRFVV ASLLENVQGR LKAWRFTDGK
301
     WQETELPRLP SGALEMTDQP WGGDVVYLAA SDFTTPLTLF ALDLNVMELT
351
401
     VMRRQPQQFD SDGINVQQFW TTSADGERIP YFHVGKNAAP DMPTLVYAYG
     GFGIPELPHY LGSIGKYWLE EGNAFVLANI RGGGEFGPRW HQAAQGISKH
451
     KSVDDLLAVV SDLSERGISS PEHIGLOGGS NGGLITAAAF VREPQSIGAL
501
VCEVPLTDMI RYPLLSAGSS WTDEYGNPQK YEVCKRRLGE LSPYHNLSDG
601 IDYPPALITT SLSDDRVHPA HALKFYAKLR ETSPQSWLYS PDGGGHTGNG
651 TQREAADELA CVLLFLKEFL G*
```

### **a041-1/m041-1** 97.9% identity in 671 aa overlap

a041-1.pep	10	20	30	40	50	60
ao41-1.pep	MKSYPDPYRHFEN	LDSAETONFA	AEANAETRAR:	FLNNDKARAL:	SDGILAQLQD	TROIPFC
m041-1	111111111111	111111111	1   1   1   1   1   1   1	11.11111		
MO41-1	MINSTEDETRICEN	PDSWELONEW	AEANAETRAR!	FLENDKARALS	DGILAOLOD	TROIPEC
	10	20	30	40	50	60
						00
	70	80	90	100	110	120
a041-1.pep	QEHRARMYHFHQD	AEYPKGVYRV	CTAATYRSGY	PEWKII EGUAT	EDET I CDDI	VT 0011011
	, , , , , , , , , , , , , , , , , , , ,					
m041-1	QEHRARMYHFHQDA	AEYPKGVYRV	CTAATYRSCYI	OFWETT FOUR		!!!!!!!
	70	80	90	100	110	
		• •	50	100	110	120
	130	140	150	1.00		
a041-1.pep			130	160	170	180
- Trop	LVEQPNRALLTLSF	COGGDIAII	CVDLEAGELVE	GGFHFPAGKN	HVSWRDENS	/WVCPAW
m041-1	LVEODNDALLTICK	(T.C.C.D.T.T.T.T.T.T.T.T.T.T.T.T.T.T.T.T.	!!!!!!!!!!!!!!	111111111	1111111	
MO41 1	DARGEMENT TO	/TRALITIES OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE LANGE OF THE	SVDLEAGELVE	GGFHFPAGKN	HVSWRDENS	WVCPAW
	130	140	150	160	170	180
	190	200	210	220	230	240
a041-1.pep	DERQLTESGYPREV	WLVERGKSFE	ESLPVYQIAE	DGMMVNAWRY	I DDAGGDED	
	• ! ! ! ! ! •					
m041-1	NERQLTQSGYPREV	WLVERGKSFE	ESLPVYOIGE	DGMMVNAWRY	1.DDOGGDTD1	TENCE
	190	200	210	220	230	
				220	230	240
	250	260	270	280	000	
a041-1.pep	FYTKTYLQVSAEGE	AKPINI PNDC	DUNCYTACHT	ZOU	290	300
	111111:11:11		DVVGILAGHL	LLTLRKDWHR	ANQSYPSGAL	VAVKLN
m041-1	:        FYTKTYLBUSAFGE	VEDI VI DADO			1111111111	11111
	FYTKTYLRVSAEGE 250	260	DV VG Y LAGHL	LLTLRKDWNR	ANQSYPSGAL	VAVKLN
	250	260	270	280	290	300
	310					
a041-1.pep		320	330	340	350	360
audi-1.pep	RGELGAAQLLFAPN	ETQALESVET	TKRFVVASLLI	ENVQGRLKAWI	RETDGKWOET	
m041-1						
modI-I	· · · · · · · · · · · · · · · · · · ·	ETQALESVET	TKRFVVASLLI	ENVOGRLKAWE	FADGKWOFU	riiii Fî Do t n
	310	320	330	340	350	
			_	0.0	550	360

نو پار از تو پار از

a041-1.pep m041-1	370 SGALEMTDQPWGGDVV !!!!!!!!!!!!! SGALEMTDQPWGGDVV 370		1111111111	1111111111	1111111111	1111
a041-1.pep m041-1	430 TTSADGERIPYFHVGK              TTSADGERIPYFHVGK 430	1111111111	1111111111	1111111111	1111111111	1111
a041-1.pep	490 RGGGEFGPRWHQAAQG               RGGGEFGPRWHQAAQG 490	1111111111	1111 11111	1111111111	1111111111	1111
a041-1.pep	550 VREPQSIGALVCEVPL	1111111111	111111111	HIIIIIIIII	HHHHH	
a041-1.pep m041-1	610 IDYPPALITTSLSDDR               IDYPPALITTSLSDDR 610	1111111111	111111111111	11111111111	11111111:1	1111
a041-1.pep	670 CVLLFLKEFLGX           CVLLFLKEFLGX 670					

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 165>: g042.seq

-					
1	ATGACGATGA	TTTGCTTGCG	CTTCCAagcG	TTCGTGCCGC	ATACCAGCGC
51	GTTATCCAAC	ACTTCCACGG	CAGCCGGCCC	TTCCTGCCCG	ATGGCGGCGG
101	TGCGGTCGAT	GATGAAAATC	CAGCCGGGGT	TTTTCTCTTT	GATGTATTCG
151			CCCTTCGTTG		
201			GTATCCAACT		
251			CCTGTAACCG		
301	TTGCCTTTGG	cggCTTCGCG	CTTTTGGGCG	AACAGCGCGT	CAATCTGCGC
351	ATTCAATTCC	GCCACGCGCG	CTTCCTTACC	GAAAATCCGC	GACAGGGTCT
401	CCATCTGCTT	CTCGCCGCTG	GTGCGGATAT	TGCCGTTGTC	CACCGTCAAA
451	TCTATGgtgG	TCGCGTTTTT	CGCCAACTGT	TCATACGCTT	CCGCACCCGG
501	CCCGCCGGTA	ATGACAAACT	GCGGATTGTG	GCGGTGCAGG	GATTCGCAAT
551	CGGGCTCAAA	CAGCGTCCCC	ACCGTTGCCG	CCTTGTCAAA	TGCAGGCTGC
601	እ እ አጥአ <i>ር</i> :				

This corresponds to the amino acid sequence <SEQ ID 166; ORF 042.ng>: g042.pep

1	MTMICLRFQA	<b>FVPHTSALSN</b>	TSTAAGPSCP	MAAVRSMMKI	QPGFFSLMYS
51	KETGCPCPSL	RKDSSTGGRP	MSPCIQLANR	DCVPKADTLL	PVTDSTSPRP
151	SMVVAFFANC	SYASAPGPPV	MTNCGLWRCR	DSQSGSNSVP	TVAALSNAGC
201	K*				

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 167>:

BNSDOCID: <WO___9957280A2_i_>

m042.seq

```
ATGACGATGA TTTGCTTGCG CTTCCAAGCG TTCGTGCCGC GTACCAGCGC
                GTTATCCAMT ACTTCGACAG CCGcCGGCCy TTCyTGCCCG ATGGCGGCGG
            51
               TACGGTCGAT GATGAAAATC CAATCGGGGT TTTTCTCTTT GATGTATTCG
           101
           151 AAGGAAACAG GCTGCCCGTG CCCCTCGTTG CGTAAAGATT CGTCTACAGG
               CGGTAGGCCG ATGTCGCCGT GTATCCAACT TGCCAACCGC GACTGCGTGC
           201
           251
                CGAAGGCGGA CACCTTGTTG CCCGTAACCG ACAGCACCAG CCCGCGTCCT
                TTGCCTTTGG CGGCTTCGCG CGTTTGGGCG AACAGCGCGT CAATCTGCGC
           301
                CTTCAATTCC GCCGCGCGC CTTCCTTGCC GAAAATCCGC GCCAAGGTCT
           351
                CCATCTGCTT TTCGCCGCTG GTGCGGATAT TGCCGTTGTC CACCGTCAGA
                TCTATGGTGG TCGCGTTTTT CGCTAACTGT TCATACGCTT CCGCGCCCGG
           451
           501
                CCCGCCGGTA ATGACAAGCT GAGGATTGTA GCGGTGCAGG GCTTCGTAAT
                CGGGCTCGAA CAGCGTCCCC ACCGTTGCCG CCTTGTCAAA TGCAGGCTGC
           551
           601 AAATAA
This corresponds to the amino acid sequence <SEQ ID 168; ORF 042>:
     m042.pep
               MTMICLRFQA FVPRTSALSX TSTAAGXSCP MAAVRSMMKI QSGFFSLMYS
            51
               KETGCPCPSL RKDSSTGGRP MSPCIQLANR DCVPKADTLL PVTDSTSPRP
               LPLAASRVWA NSASICAFNS AARASLPKIR AKVSICFSPL VRILPLSTVR
           101
           151
               SMVVAFFANC SYASAPGPPV MTSXGLXRCR ASXSGSNSVP TVAALSNAGC
           201
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 169>:
     a042.seq
               ATGACGATGA TTTGCTTGCG CTTCCAAGCG TTCGTGCCGC GTACCAGCGC
           51
               GTTATCCAAT ACTTCGACAG CCGCCGGCCC TTCCTGCCCG ATGGCGGCGG
          101
               TACGGTCGAT GATGAAAATC CAATCGGGGT TTTTCTCTTT GATGTATTCG
               AAGGAAACAG GCTGCCCGTG CCCCTCGTTG CGTAAAGATT CGTCTACAGG
          151
               CGGTAGGCCG ATGTCGCCGT GTATCCAACT TGCCAACCGC GACTGCGTGC
          201
          251
               CGAAGGCGGA CACCTTGTTG CCCGTAACCG ACAGCACCAG CCCGCGTCCT
          301
               TTGCCTTTGG CGGCTTCGCG CGTTTGGGCG AACAGCGCGT CAATCTGCGC
               CTTCAATTCC GCCGCGCGC CTTCCTTGCC GAAAATCCGC GCCAAGGTCT
          351
               CCATCTGCTT TTCGCCGCTG GTGCGGATAT TGCCGTTGTC CACCGTCAGA
          401
               TCTATGGTGG TCGCGTTTTT CGCCAACTGT TCATACGCTT CCGCGCCCGG
          451
               CCCGCCGGTA ATGACAAGCT GAGGATTGTA GCGGTGCAGG GCTTCGTAAT
               CGGGCTCGAA CAGCGTCCCC ACCGTTGCCG CCTTGTCAAA TGCAGGCTGC
          551
          601
               AAATAA
This corresponds to the amino acid sequence <SEQ ID 170; ORF 042.a>:
     a042.pep
               MTMICLRFQA FVPRTSALSN TSTAAGPSCP MAAVRSMMKI QSGFFSLMYS
           51
               KETGCPCPSL RKDSSTGGRP MSPCIQLANR DCVPKADTLL PVTDSTSPRP
               LPLAASRVWA NSASICAFNS AARASLPKIR AKVSICFSPL VRILPLSTVR
               SMVVAFFANC SYASAPGPPV MTS*GL*RCR AS*SGSNSVP TVAALSNAGC
          151
          201
               K*
m042/a042
            99.0% identity over a 201 aa overlap
                                             30
                                                       40
                  MTMICLRFQAFVPRTSALSXTSTAAGXSCPMAAVRSMMKIQSGFFSLMYSKETGCPCPSL
     m042.pep
                  MTMICLRFQAFVPRTSALSNTSTAAGPSCPMAAVRSMMKIQSGFFSLMYSKETGCPCPSL
     a042
                                  . 20
                                             30
                                                       40
                                                                50
                                                                          60
                                   80
                                             90
                                                     100
                                                               110
    m042.pep
                 RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVTDSTSPRPLPLAASRVWANSASICAFNS
                  a042
                 RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVTDSTSPRPLPLAASRVWANSASICAFNS
                         70
                                   80
                                             90
                                                     100
                                                               110
                                                                         120
                        130
                                  140
                                            150
                                                     160
                                                               170
                                                                         180
                 {\tt AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANCSYASAPGPPVMTSXGLXRCR}
    m042.pep
```

1

AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANCSYASAPGPPVMTSXGLXRCR a042 130 140 150 170 190 200 ASXSGSNSVPTVAALSNAGCKX m042.pep a042 ASXSGSNSVPTVAALSNAGCKX 190 200

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae

ORF 042 shows 93.0% identity over a 201 aa overlap with a predicted ORF (ORF 042.ng) from N. gonorrhoeae:

m042/g042

	10	20	30	40	50	60
m042.pep	MTMICLRFQAFVPR	TSALSXTSTA	AGXSCPMAAV	RSMMKIQSGE	FSLMYSKETG	CPCPSL
	11111111111111111	11111 1111	11 111111	11111111111	11111111111	111111
g042	MTMICLRFQAFVPH	TSALSNTSTA	AGPSCPMAAV	RSMMKIQPGE	FSLMYSKETG	CPCPSL
	10	20	30	40	50	60
	70	80	90	100	110	120
m042.pep	RKDSSTGGRPMSPC	IQLANRDCVP	KADTLLPVTD	STSPRPLPLA	ASRVWANSAS	ICAFNS
			1111111111	1111111111	111 11111	11111
g042	RKDSSTGGRPMSPC	IQLANRDCVP	KADTLLPVTD	STSPRPLPLA	ASRFWANSAS	ICAFNS
	70	80	90	100	110	120
	130	140	150	160	170	180
m042.pep	AARASLPKIRAKVS	ICFSPLVRIL	PLSTVRSMVV.	AFFANCSYAS	APGPPVMTSX	GLXRCR
	1:1111111 :11			1131111111	11111111:	11 111
g042	ATRASLPKIRDRVS	ICFSPLVRIL	PLSTVKSMVV.	AFFANCSYAS	APGPPVMTNC	GLWRCR
	130	140	150	160	170	180
	190	200				
m042.pep	ASXSGSNSVPTVAA	LSNAGCKX				
	1 1111111111					
g042	DSQSGSNSVPTVAA					
	190	200				

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 171>: m042-1.seq

•	1.50	4				
	1	ATGACGATGA	TTTGCTTGCG	CTTCCAAGCG	TTCGTGCCGC	GTACCAGCGC
	51	GTTATCCAAT	ACTTCGACAG	CCGCCGGCCC	TTCCTGCCCG	ATGGCGGCGG
	101	TACGGTCGAT	GATGAAAATC	CAATCGGGGT	TTTTCTCTTT	GATGTATTCG
	151	AAGGAAACAG	GCTGCCCGTG	CCCCTCGTTG	CGTAAAGATT	CGTCTACAGG
	201	CGGTAGGCCG	ATGTCGCCGT	GTATCCAACT	TGCCAACCGC	GACTGCGTGC
	251		CACCTTGTTG			
	301	TTGCCTTTGG	CGGCTTCGCG	CGTTTGGGCG	AACAGCGCGT	CAATCTGCGC
	351	CTTCAATTCC	GCCGCGCGCG	CTTCCTTGCC	GAAAATCCGC	GCCAAGGTCT
	401	CCATCTGCTT	TTCGCCGCTG	GTGCGGATAT	TGCCGTTGTC	CACCGTCAGA
	451	TCTATGGTGG	TCGCGTTTTT	CGCTAACTGT	TCATACGCTT	CCGCGCCCGG
	501	CCCGCCGGTA	A			

This corresponds to the amino acid sequence <SEQ ID 172; ORF 042-1>:

m042-1.pep 1 MTMICLRFQA FVPRTSALSN TSTAAGPSCP MAAVRSMMKI QSGFFSLMYS

51 KETGCPCPSL RKDSSTGGRP MSPCIQLANR DCVPKADTLL PVTDSTSPRP 101 LPLAASRVWA NSASICAFNS AARASLPKIR AK<u>VSICFSPL VRILPLSTV</u>R 151 SMVVAFFANC SYASAPGPPV MTS*

m042-1/g042 95.4% identity in 173 aa overlap

BNSDOCID: <WO___9957280A2_j_>

m042-1.pep	111111111	AFVPRTSALSNTS    :        AFVPHTSALSNTS	30 TAAGPSCPMAAVF           TAAGPSCPMAAVF 30		11111111	LIBILI
m042-1.pep	111111111	MSPCIQLANRDC	90 VPKADTLLPVTDS             VPKADTLLPVTDS 90	100 STSPRPLPLAA	110 SRVWANSAS	120 ICAFNS
m042-1.pep	1:1111111	AKVSICFSPLVR	150 ILPLSTVRSMVVA           ILPLSTVKSMVVA 150		1111111	SLWRCR 180
g042	DSQSGSNSVP 190	TVAALSNAGCKX 200				
The following pa	rtial DNA sequ	ence was iden	ntified in N. me	eningitidis <	SEQ ID 1	73>:
101 1201 1201 1201 1201 1201 1201 1201	ATGACGATGA TT. GTTATCCAAT ACT. FACGGTCGAT GAT. AAGGAAACAG GCT. CGAAGGCGGA CACT. CTATCAATTCC GCC. CCATCTGCTT TT. CCTATGGTGG TCC. CCCGCCGGTA A  to the amino act. CTMICLRFQA FVE. CETGCPCPSL RKD. PLAASRVWA NSA	TTCGACAG CCGC TGAAAATC CAAT TGCCCGTG CCCC GTCGCCGT GTAT CCTTGTTG CCCG CCTTCGCG CGTTC CGCCGCTG GTGC CGCCGCTT CGCC CGCTTTTT CGCC  CTTCGCC CTTC CGCCGCTG GTGC CGCTTTTT CGCC  CTTCGCC  CTTC	CCGGCCC TTCCTCCGGGGT TTTTCCTCGGGGT TTTTCCTCGTTG CGTAACCCAACT TGCCAACTCTGGCG AACAGCCTTGCC GAAAATGGATAT TGCCGTAACTGT TCATACCGAACTGT TCATACCAACTGAACTG	GCCCG ATGGC FCTTT GATGT AGATT CGTCT ACCGC GACTG ACCAG CCCGC CGCGT CAATC FCCGC GCCAA FTGTC CACCG CGCTT CCGCG  ORF 042-1.	CGGCGG CATTCG CACAGG CGTGC CGTCCT CTGCGC GGTCT CCAGA CCCGG	
m042-1/a042	-1 100.0%	identity in	173 aa overla	p		
m042-1.pep	1 1 1 1 1 1 1 1 1 1	1   1   1   1   1   1   1   1   1   1	30 AAGPSCPMAAVRS            AAGPSCPMAAVRS 30			1111
m042-1.pep			90 PKADTLLPVTDST           PKADTLLPVTDST 90	11111111111		1.1.1.1
m042-1.pep	1 1 1 1 1 1 1 1 1 1 1 1		150 .PLSTVRSMVVAF             PLSTVRSMVVAF   150	11111111111	111111	

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 175>:

g043.seq					
1	ATGGTTGTTT	CAAATCAAAA	TATCTATGCC	GTCGGCCCAT	CAGCACTTTT
51	TCACATCCGA	AGGCAAAAAT	CCGTAATGCC	GCCTGAACGC	TTCgttgaAC
101	CGTCCCGCGT	ggcggtagcc	gcAAAAGTGC	ATCGCGGCTT	GGATGGTGCT
151	GCCCGATTCG	ATGAGGGcga	gcGCGTGTTC	CAGCCGCAGG	CGGCGCAGGC
201	GTCCGGCGAC	GGTTTCGCCG	GTTTGCGCTT	TGAAATAGCG	TTTCAGGTAG
251	CATTCGTTCA	GCCCGACGCG	GCGGGCGATT	TCGGCGATGG	TCAGCGGGCG
301	GGCGAATTCG	CTGTTCAAAA	TATCGGCGGC	TTCGTCTATG	CGCCGGCGGC
351	GGTAGCCGTT	GTCGTGGCGG	CGGAAGGTGA	AGCGTAA	

### This corresponds to the amino acid sequence <SEQ ID 176; ORF 043.ng>:

g043.pep

- 1 MVVSNQNIYA VGPSALFHIR RQKSVMPPER FVEPSRVAVA AKVHRGLDGA 51 ARFDEGERVF QPQAAQASGD GFAGLRFEIA FQVAFVQPDA AGDFGDGQRA
- 101 GEFAVONIGG FVYAPAAVAV VVAAEGEA*

#### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 177>:

m043.seq

1	ATGGTTGTTT	CAAATCAAAA	TATCTATGCC	GCCGGCCCCT	CAGCACTTCT
51	TCACATCCGA	AGGCAAAAAT	CCGTAATGCC	GTCTGAACGC	TTCGTTGAAC
101	CGTCCCGCGT	GGCGGTAGCC	GCAAAAGTGC	ATGGCGGCTT	GGACGGTGCT
151	GCCGGATTCG	ATGAGGGCGA	GCGCGTGTTC	CAGCCGCAGG	CGGCGCAgGC
201	ATCCGGCGAC	GGTTTCGCCG	GTTTGCGCTT	TGAAATAGCG	TTTCAGGTAG
251	CATTCGTTCA	GTCCGACGCG	GCGGGCGATT	TCGGCGATGG	TCAGCGGACG
301	GGCGAATTCG	TGTTGCAGGA	TGTCGGCGGC	TTCGTCTATG	CGCCGACGGC
351	GGTAACCGTT	GTCGTGGCGG	CGGAAGGTGA	AGCGCAATAA	

#### This corresponds to the amino acid sequence <SEQ ID 178; ORF 043>:

m043.pep

- 1 MVVSNQNIYA AGPSALLHIR RQKSVMPSER FVEPSRVAVA AKVHGGLDGA
  - 51 AGFDEGERVF QPQAAQASGD GFAGLRFEIA FQVAFVQSDA AGDFGDGQRT 101 GEFVLQDVGG FVYAPTAVTV VVAAEGEAQ*

Computer analysis of the amino acid sequences gave the following results: Homology with a predicted ORF from N. meningitidis menA with menB

ORF 043 shows 89.8% identity over a 128 aa overlap with a predicted ORF (ORF043.a) from N. gonorrhoeae:

#### m043/g043

	10	20	30	40	50	60
m043.pep	MVVSNQNIYAAGPS	ALLHIRRQKS	VMPSERFVE	SRVAVAAKVE	IGGLDGAAGFI	DEGERVF
		11:1111111	111 111111		- 111111 11	111111
g043	MVVSNQNIYAVGPS	ALFHIRRQKS	VMPPERFVE	PSRVAVAAKVE	IRGLDGAARFI	DEGERVF
	10	20	30	40	50	60
	70	80	90	100	110	120
m043.pep	QPQAAQASGDGFAG	LRFEIAFQVA				
	11111111111111	1111111111	111 111111		:1::11111	1:11:1
g043	QPQAAQASGDGFAG	LRFEIAFQVA	FVQPDAAGDI	GDGQRAGEFA	VQNIGGFVY#	APAAVAV
	70	80	90	100	110	120
	130					
m043.pep	VVAAEGEAQX					
	1111111					
g043	VVAAEGEAXX					
	130					

#### The following partial DNA sequence was identified in N. meningitidis <SEO ID 179>:

a043.seq

- 1 ATGGTTGTTT CAAATCAAAA TATCTATGCC GCCGGCCCCT CAGCACTTCT 51 TCACATCCGA AGGCAAAAAT CCGTAATGCC GTCTGAACGC TTCGTTGAAC
- 101 CGTCCCGCGT GGCGGTAGCC GCAAAAGTGC ATGGCGGCTT GGACGGTGCT



251 C 301 G	ATCCGGCGAC CATTCGTTCA GGCGAATTCG	GTCCGACGCG	GTTTGCGCTT GCGGGCGATT TGTCGGCGGC	TGAAATAGCG TCGGCGATGG TTCGTCTATG	CGGCGCAGGC TTTCAGGTAG TCAGCGGACG CGCCGACGGC
----------------	----------------------------------------	------------	----------------------------------------	----------------------------------------	------------------------------------------------------

### This corresponds to the amino acid sequence <SEQ ID 180; ORF 043.a>:

```
a043.pep

1 MVVSNQNIYA AGPSALLHIR RQKSVMPSER FVEPSRVAVA AKVHGGLDGA
51 AGFDEGERVF QPQAAQASGD GFAGLRFEIA FQVAFVQSDA AGDFGDGQRT
101 GEFVLQDVGG FVYAPTAVTV VVAAEGEAQ*
```

m043/a043	100.0%	identity	in	129	aa	overlap
				147	aa	Overrap

m043.pep	10 MVVSNQNIYAAGPS            MVVSNQNIYAAGPS 10	1   1   1   1   1   1   1	11111111	111111111	1111111111	111111
m043.pep	70 QPQAAQASGDGFAG            QPQAAQASGDGFAG 70	111111111	1	111111111	1111111111	LELLI
m043.pep	130 VVAAEGEAQX           VVAAEGEAQX 130					

# The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 181>: g044.seq

```
1 ATGCTGCCCG ACCAGAGCGT CGAGTTCTTG CCACAAGTCG TCGTTTTTGA
51 CGGGCTGTTT GGCGGCGGTT TTCCAGCCGT TGCGCTTCCA ACCGTGTATC
101 CAGTTTTCCA TGCCGTTTTT GACGTATTGC GAGTCGGTGC AGATGATGAC
151 GGTGCAGCGG CGTTTGAGCG ATTTCAGCCC TTCGATAACG GCGGTCAGCT
201 CCATGCGGTT GTTGGTGGTT TGCGCTTCGC CGCCGAAAAG TTCTTTTTCG
251 CGGCTGCCGT AGCGCATTAA
```

# This corresponds to the amino acid sequence <SEQ ID 182; ORF 044.ng>: g044.pep

1 MLPDQSVEFL PQVVVFDGLF GGGFPAVALP TVYPVFHAVF DVLRVGADDD 51 GAAAFERFQP FDNGGQLHAV VGGLRFAAEK FFFAAAVAH*

## The following partial DNA sequence was identified in N. meningitidis <SEQ ID 183>: m044.seq

1	ATGCCGTCCG	ACTAGAGCGT	CGAGTTCTTT	CCAGAAGTCG	TCGTTTTTCA
21	CGGGCTGTTT	GGAGGCGGTT	TTCCAGCCGT	TGCGCTTCCA	ACCGTGTATC
101	CAGTTTTCCA	TGCCATTTTT	GACGTATTGC	GAGTCGGTGC	ACATCATCAC
151	GGTGCAGCGG	CGTTTGAGCG	ATTTCAGTCC	TTCGATGACG	GCAGTCAGTT
201	CCATGCGGTT	GTTGGTGGTT	TGCGCTTCGC	CGCCGAAAAG	TTCTTTTTCC
251	TGGCTACCGT	AGCGCAyTAa		CCCCAAAAG	1101111100

This corresponds to the amino acid sequence <SEQ ID 184; ORF 044>: m044.pep

- 1 MPSDXSVEFF PE<u>VVVFDGLF GGGFPAVAL</u>P TVYPVFHAIF DVLRVGADDD
- 51 GAAAFERFOS FDDGSQFHAV VGGLRFAAEK FFFVATVAH*

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 185>:

```
a044.seq

1 GTGCCGTCCG ACCAGCGGT CGAGTTCTTT CCACAAGTCG TCGTTTTTGA
51 CGGGCTGTTT GGCGGCGGT TTCCAGCCGT TGCGCTTCCA ACCGTGTATC
101 CAGTTTTCCA TGCCGTTTTT GACGTATTGC GAGTCGGTGC AGATGATGAC
151 GGTGCAGCGG CGTTTGAGCG ATTTCAGTCC TTCGATGACG GCGGTCAGTT
201 CCATACGGTT GTTGGTGGTT TGCGCTTCGC CGCCGAAAAG TTCTTTTTCG
251 TGGCTGCCGT AGCGCATTAA
```

This corresponds to the amino acid sequence <SEQ ID 186; ORF 044.a>:

a044.pep

1 VPSDQRVEFF PQVVVFDGLF GGGFPAVALP TVYPVFHAVF DVLRVGADDD

51 GAAAFERFQS FDDGGQFHTV VGGLRFAAEK FFFVAAVAH*

m044/a044 91.0% identity over a 89 aa overlap

```
10
                                  30
                                          40
                                                  50
          MPSDXSVEFFPEVVVFDGLFGGGFPAVALPTVYPVFHAIFDVLRVGADDDGAAAFERFQS
m044.pep
               a044
          VPSDQRVEFFPQVVVFDGLFGGGFPAVALPTVYPVFHAVFDVLRVGADDDGAAAFERFOS
                 10
                         20
                                  30
                                          40
                 70
                         80
          FDDGSQFHAVVGGLRFAAEKFFFVATVAHX
m044.pep
           1111:111:111:1111
a044
          FDDGGQFHTVVGGLRFAAEKFFFVAAVAHX
                 70
                         80
                                  90
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae

ORF 044 shows 86.5% identity over a 89 aa overlap with a predicted ORF (ORF 044.ng) from N. gonorrhoeae:

m044/g044

```
30
                                        40
                                                50
                                                        60
m044.pep
          MPSDXSVEFFPEVVVFDGLFGGGFPAVALPTVYPVFHAIFDVLRVGADDDGAAAFERFOS
            {\tt MLPDQSVEFLPQVVVFDGLFGGGFPAVALPTVYPVFHAVFDVLRVGADDDGAAAFERFQP}
q044
                10
                        20
                                30
                                        40
                                                        60
                70
                        80
                                90
m044.pep
          FDDGSQFHAVVGGLRFAAEKFFFVATVAHX
          g044
          FDNGGQLHAVVGGLRFAAEKFFFAAAVAHX
                70
                        80
                                90
```

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 187>:

```
ATGTCGGCAA TGCTGCGTCC GACAAGCAGC CCGCCGCgcc gCGCCTGTAT
 1
 51
    GATGACCATC CGCACGCGT CGTCTGCAAA ACGTAAAACC TGCAATGCGC
    CCGGGCAGTC TATCAGGCCG GCAAGCTGTT CGGTAACGAG CTGTTCGGGG
101
    CTGATGGTTT CGGTTATGCC gaATATGGAA AGGCTGCCGt TTTcGTTGTT
151
201
    TTCGAGCTTG GGGCTGAGGT ATTCGAGGTA TtcgctGGAA CGGACGCGCG
251
    CGATGCGGCC GGGGATGTTG AACAGGTCGG CGGCAACTTT GCAGGCGACG
    ATGTTGGTTT CGTCGCTGCG GGagaGCGCG AGcagcaagt cggcatcttC
301
    CgcgccggcG Cgttataatg tgAAGGGGGA TGCGccgttg ccgaAAACGG
351
401
    TTTGGacatc gaggcggctg CCTGTTTCCT GCAATGCTTT TTCGTCGATG
451
    TCGATAAcgg TTACGTCGTT GTTGGTGATG GCGGCAAGGT TTTGCGCGAC
```

BNSDOCID: <WO___9957280A2_i_>

WO 99/57280



```
501 GGTAGAACCT ACCTGCCCGT TGCCTAAAAT GAGGATTTTC ACGGTATGGG
            551 TCGCCGGGTG A
 This corresponds to the amino acid sequence <SEQ ID 188; ORF 046.ng>:
      g046.pep
                MSAMLRPTSS PPRRACMMTI RTRSSAKRKT CNAPGQSIRP ASCSVTSCSG
             1
            51 LMVSVMPNME RLPFSLFSSL GLRYSRYSLE RTRAMRPGML NRSAATLQAT
           101 MLVSSLRESA SSKSASSAPA RYNVKGDAPL PKTVWTSRRL PVSCNAFSSM
           151 SITVTSLLVM AARFCATVEP TCPLPKMRIF TVWVAG*
 The following partial DNA sequence was identified in N. meningitidis <SEQ ID 189>:
      m046.seq
             1 ATGTCGGCAA TGCTGCGTCC GACAAGCAST CCGC.r.sGC gCGcCTGTAT
            51 GATGACCATC CGCACGCGGT CGTCTGCAAA ACGTAAAACC TGCAATGCGC
           101 CCGGGCAGTC TATCAGGCCG GCAAGCTGTT CGGTAACGAG CTGTTCGGGG
           151 CTGATGGTTT CGGTTATGCC GAATATGGAA AGGCTGCCGT TTTCGTTGTT
                TTCGAGCTTG GGGCTGAGGT ATTCGAGGTA TTCGCTGGAA CGGACGCGCG
           251 CGATGCGGCC GGGGATGTTG AACAGGTCGG CGGCAACTTT GCAGGCGACG
           301 ATGTTGGTTT CGTCGCTGCG GGAGAGCGCG AGCAGCAAGT CGGCATCTTC
           351 CGCGCCGGCG CGTTCTAATG TGAAGGGGGA TGCGCCGTTG CCGAAAACGG
           401 TTTGGACATC GAGGCGGCTG CCTGTTTCCT GCAATGCTTT TTCGTCGATG
           451 TCGATAACGG TTACGTCGTT GTTGGGTATG GCGGCAAGGT TTTGTGCGAC
           501 GGTAGAACCT ACCTGTCCGT TGCCTAAAAT GAGGATTTTC ACGGTGTGGG
           551 TCGCCGAGTG A
This corresponds to the amino acid sequence <SEQ ID 190; ORF 046>:
      m046.pep
               MSAMLRPTSX PXXRACMMTI RTRSSAKRKT CNAPGQSIRP ASCSVTSCSG
           51 LMVSVMPNME RLPFSLFSSL GLRYSRYSLE RTRAMRPGML NRSAATLQAT
          101 MLVSSLRESA SSKSASSAPA RSNVKGDAPL PKTVWTSRRL PVSCNAFSSM
          151 SITVTSLLGM AARFCATVEP TCPLPKMRIF TVWVAE*
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 191>:
     a046.seq
               ATGTCGGCAA TGCTGCGTCC GACAAGCAGT CCGCCGCGCC GCGCCTGTAT
            1
               GATGACCATC CGCACGCGGT CGTCTGCAAA ACGTAAAACC TGCAATGCGC
           51
          101
               CCGGGCAGTC TATCAGGCCG GCAAGCTGTT CGGTAACGAG CTGTTCGGGG
               CTGATGGTTT CGGTTATGCC GAATATGGAA AGGCTGCCGT TTTCGTTGTT
          151
          201 TTCGAGCTTG GGGCTGAGGT ATTCGAGGTA TTCGCTGGAA CGGACGCGCG
              CGATGCGGCC GGGGATGTTG AACAGGTCGG CGGCAACTTT GCAGGCGACG
          301 ATGTTGGTTT CGTCGCTGCG GGAGAGCGCG AGCAGCAAGT CGGCATCTTC
          351
               CGCGCCGGCG CGTTCTAATG TGAAGGGGGA TGCGCCGTTG CCGAAAACGG
          401
               TTTGGACATC GAGGCGGCTG CCTGTTTCCT GCAATGCTTT TTCGTCGATG
          451 TCGATAACGG TTACGTCGTT GTTGGGTATG GCGGCAAGGT TTTGTGCGAC
               GGTAGAACCT ACCTGTCCGT TGCCTAAAAT GAGGATTTTC ACGGTGTGGG
          551 TCGCCGAGTG A
This corresponds to the amino acid sequence <SEQ ID 192; ORF 046.a>:
     a046.pep
              MSAMLRPTSS PPRRACMMTI RTRSSAKRKT CNAPGQSIRP ASCSVTSCSG
            1
               LMVSVMPNME RLPFSLFSSL GLRYSRYSLE RTRAMRPGML NRSAATLQAT
          101 MLVSSLRESA SSKSASSAPA RSNVKGDAPL PKTVWTSRRL PVSCNAFSSM
          151 SITVTSLLGM AARFCATVEP TCPLPKMRIF TVWVAE*
            98.4% identity over a 186 aa overlap
m046/a046
                                             30
                                                       40
                 MSAMLRPTSXPXXRACMMTIRTRSSAKRKTCNAPGQSIRPASCSVTSCSGLMVSVMPNME
     m046.pep
                  a046
                 MSAMLRPTSSPPRRACMMTIRTRSSAKRKTCNAPGQSIRPASCSVTSCSGLMVSVMPNME
```

	10	20	30	40	50	60
	70	80	90	100	110	120
m046.pep	RLPFSLFSSLGI	RYSRYSLERT	RAMRPGMLNRSA	ATLQATMLV:	SSLRESASSKS	SASSAPA
		111111111	111111111111	111111111		
a046	RLPFSLFSSLGI	RYSRYSLERT	RAMRPGMLNRSA	ATLQATMLV:	SSLRESASSKS	SASSAPA
	70	80	90	100	110	120
	130	140	150	160	170	180
m046.pep	RSNVKGDAPLPK	TVWTSRRLPV	SCNAFSSMSITV	TSLLGMAARI	FCATVEPTCPI	PKMRIF
		111111111	111111111111	1111111	[	
a046	RSNVKGDAPLPK	TVWTSRRLPV	SCNAFSSMSITV	TSLLGMAARI	FCATVEPTCPI	PKMRIF
	130	140	150	160	170	180
m046.pep	TVWVAEX					
0.1.6	111111					
a046	TVWVAEX				•	

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae
ORF 046 shows 97.3% identity over a 185 aa overlap with a predicted ORF (ORF 046.ng) from N. gonorrhoeae:

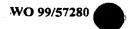
m046/g046

	10	20	30	40	50	60
m046.pep	MSAMLRPTSXPXXI	RACMMTIRTRS	SAKRKTCNAP	GQSIRPASCS	VTSCSGLMVS	VMPNME
			111111111		11111111	11111
g046	MSAMLRPTSSPPRI		SAKRKTCNAP	GOSIRPASCS	VTSCSGLMVS	VMPNME
	10	20	30	40	50	60
	70	80	90	100	110	120
m046.pep	RLPFSLFSSLGLR	SRYSLERTRA	MRPGMLNRSA	ATLOATMLVS	SLRESASSKS	ASSAPA
			11111111	11111111111		
g046	RLPFSLFSSLGLR	(SRYSLERTRA	MRPGMLNRSA	ATLQATMLVS	SLRESASSKS	ASSAPA
	70	80	90	100	110	120
	130	140	150	160	170	180
m046.pep	RSNVKGDAPLPKT	WTSRRLPVSC	NAFSSMSITV	TSLLGMAARF	CATVEPTCPL	PKMRIF
				1111111	1 1 1 1 1 1 1 1 1 1	111111
g046	RYNVKGDAPLPKTV	WTSRRLPVSC	NAFSSMSITV	TSLLVMAARF	CATVEPTCPL	PKMRIF
	130	140	150	160	170	180
m046.pep	TVWVAEX					
g046	TVWVAGX					

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 193>: g047.seq

1	ATGGTCATCA	TACAGGCGcg	gcGCGGCGGG	CTGCTTGTCG	GACGCAGCAT
51	TGCCGACATC	GCCCAAGATT	TGCCCGACGG	GGCCGACTGC	CAAATCTGCG
101	CCGTTTACCG	CAACAACCGC	CTCATCGTCC	CCGCGCCGCA	AACCGTCATC
151	ATCGAAGGCG	ACGAAATCCT	GTTTGCCGCC	GCCGCCGAAA	ACATCGGGGC
201	GGTCATACCC	gaATTGCGCC	CCAAAGAAAC	CAGCACCCGC	CGCATCATGA
251	TTGCCGGCGG	CGGCAACATC	tgctACCGCC	TCGCCAAGCA	GCTCGAACAC

BNSDOCID: <WO___9957280A2_l_>



301	GCATAcaacG	TCAAAATCAT	CGAATGCCGG	CCGCGCcata	ccgaATGGAT
351	AGCCGAAAAC	ctcgAcaaCA	CCCTCGTCCT	GCAAGGTTCG	Gcaaccgacg
401	aAaccctgct	cgAcaacgaa	tacatcgacg	aaatcgaCGT	ATTCTCCCC
451	CTGACCAACG	ACGACGAAAG	CAACATTAtq	tCCGCCCTTT	TGGCGAAAAA
501	CCTcggcgCG	AAGCgcgtca	tcgqCATCGT	CAACCGCTCA	AGCTACGTCG
551	ATTTGCTCGA	AGGCAACAAA	ATCGACATCG	TCGTCTCCCC	CCACCTCATC
601	ACCATCGGCT	CGATACTCGC	CCACATCCGG	CGCGGCGACA	TCGTTGCCGT
651	CCACCCCATC	CGGCGCGCA	CGGCGGAAGC	CATCGAAGTC	GTCGCGCNCG
701	GCGACAAAAA	AACTTCCGCC	ATCATCGGCA	GGCGCATCAG	CGGCATCAAA
751	TGGCCCGAAG	GCTGCCACAT	TGCCGCCGTC	GTCCGCGCCG	GAACCGGCGA
801	AACCATTATG	GGACACCATA	CCGAAACCGT	CATCCAAGAC	GGTGACCACA
851	TCATCTTTTT	CGTCTCGCGC	CGGCGCATCC	TGAACGAACT	GGAGAAACTC
901	ATCCAAGTCA	AAATGGGCTT	TTTCGCATAA		

## This corresponds to the amino acid sequence <SEQ ID 194; ORF 047.ng>:

```
1 MVIIQARRGG LLVGRSIADI AQDLPDGADC QICAVYRNNR LIVPAPQTVI
51 IEGDEILFAA AAENIGAVIP ELRPKETSTR RIMIAGGGNI CYRLAKQLEH
101 AYNVKIIECR PRRAEWIAEN LDNTLVLQGS ATDETLLDNE YIDEIDVFCA
151 LTNDDESNIM SALLAKNLGA KRVIGIVNRS SYVDLLEGNK IDIVVSPHLI
201 TIGSILAHIR RGDIVAVHPI RRGTAEAIEV VAHGDKKTSA IIGRRISGIK
251 WPEGCHIAAV VRAGTGETIM GHHTETVIQD GDHIIFFVSR RRILNELEKL
301 IOVKMGFFG*
```

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 195>: m047.seq

```
ATGGTCATCA TACAGGCGCG C..syGCGGA sTGCTTGTCG GACGCAGCAT
     TGCCGACATC GCCCAAGATT TGCCCGACGG GGCCGACTGC CAAATCTGCG
101 CCGTTTACCG CAACAACCGC CTCATCGTCC CCGCGCCGCA AACCGTCATC
151 ATCGAAGGCG ACGAAATCCT ATTTGCCGCC GCCGCCGAAA ACATCGGCGC
201 GGTCATACCC GAATTGCGCC CCAAAGAAAC CCAAAGAAAC CAGCcCmgmm
251 GCATCATGAT TKCCGGCGGC GGCAACATCG GCTACCGTCT CGCCAAGCAG
301 CTCGAACACG CATACAACGT YAAAATCATC GAATGCCGGC CGCGCCGTGC
351 CGAATGGATA GCCGAAAACC TCGACAACAC CCTCGTCYTG CAAGGTTCGG
401 CAACCGACGA AACCCTGCTC GACAACGAAT ACATCGACGA AATCGACGTA
451 TTCTGCGCCC TGACCAACGA CGACGAAAGC AACATTATGT CCGCCCTTTT
501 GGCGAAAAAC CTCGGCGCGA AGCGCGTCAT CGGCATCGTC AACCGCTCAA
551 GCTACGTCGA TTTGCTCGAA GGCAACAAAA TCGACATCGT CGTCTCCCCC
    CACCTCATCA CCATCGGCTC GATACTCGCC CACATCCGGC GCGGCGACAT
601
651 CGTTGCCGTC CACCCCATCC GGCGCGCAC GGCGGAAGCC ATCGAAGTCG
701 TCGCACACGG CGACAAAAAA ACTTCCGCCA TCATCGGCAG GCGCATCAGC
751 GGCATCAAAT GGCCCGAAGG CTGCCACATT GCCGCCGTCG TCCGCGCCGG
801 AACCGGCGAA ACCATTATGG GACACCATAC CGAAACCGTC ATCCAAGACG
851 GCGACCACAT CATCTTTTC GTCTCGCGCC GGCGCATCCT GAACGAACTG
901 GAAAAACTCA TCCAGGTCAA AATGGGCTTT TTCGGATAA
```

## This corresponds to the amino acid sequence <SEQ ID 196; ORF 047>: m047.pep

```
1 MVIIQARXXG XLVGRSIADI AQDLPDGADC QICAVYRNNR LIVPAPQTVI
51 IEGDEILFAA AAENIGAVIP ELRPKETQRN QPXXIMIXGG GNIGYRLAKQ
101 LEHAYNVKII ECRPRRAEWI AENLDNTLVL QGSATDETLL DNEYIDEIDV
151 FCALTNDDES NIMSALLAKN LGAKRVIGIV NRSSYVDLLE GNKIDIVVSP
201 HLITIGSILA HIRRGDIVAV HPIRRGTAEA IEVVAHGDKK TSAIIGRRIS
251 GIKWPEGCHI AAVVRAGTGE TIMGHHTETV IQDGDHIIFF VSRRRILNEL
301 EKLIQVKMGF FG*
```

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 197>:

ATGGTCATCA TACAGGCGCG GCGCGGCGGA CTGCTTGTCG GACGCAGCAT

11. 1 No. 1 12.1 4

51	TGCCGACATC	GCCCAAGATT	TGCCCGACGG	GGCCGACTGC	CAAATCTGCG
101	CCGTTTACCG	CAACAACCGC	CTCATCGTCC	CCGCGCCGCA	AACCGTCATC
151	ATCGAAGGCG	ACGAAATCCT	ATTTGCCGCC	GCCGCCGAAA	ACATCGGCGC
201	GGTCATACCC	GAATTGCGCC	CCAAAGAAAC	CAGCACCCGC	CGCATCATGA
251	TTGCCGGCGG	CGGCAACATC	GGCTACCGTC	TCGCCAAGCA	GCTCGAACAC
301	GCATACAACG	TCAAAATCAT	CGAATGCCGG	CCGCGCCGTG	CCGAATGGAT
351	AGCCGAAAAC	CTCGACAACA	CCCTCGTCCT	GCAAGGTTCG	GCAACCGACG
401	AAACCCTGCT	CGACAACGAA	TACATCGACG	AAATCGACGT	ATTCTGCGCC
451	CTGACCAACG	ACGACGAAAG	CAACATTATG	TCCGCCCTTT	TGGCGAAAAA
501	CCTCGGCGCG	AAGCGCGTCA	TCGGCATCGT	CAACCGCTCA	AGCTACGTCG
551	ATTTGCTCGA	AGGCAACAAA	ATCGACATCG	TCGTCTCCCC	CCACCTCATC
601	ACCATCGGCT	CGATACTCGC	CCACATCCGG	CGCGGCGACA	TCGTTGCCGT
651	CCACCCCATC	CGGCGCGCA	CGGCGGAAGC	CATCGAAGTC	GTCGCACACG
701	GCGACAAAAA	AACTTCCGCC	ATCATCGGCA	GGCGCATCAG	CGGCATCAAA
751	TGGCCCGAAG	GCTGCCACAT	TGCCGCCGTC	GTCCGCGCCG	GAACCGGCGA
801	AACCATTATG	GGACACCATA	CCGAAACCGT	CATCCAAGAC	GGCGACCACA
851	TCATCTTTTT	CGTCTCGCGC	CGGCGCATCC	TGAACGAACT	GGAAAAACTC
901	ATCCAAGTCA	AAATGGGCTT	TTTCGGATAA		

#### This corresponds to the amino acid sequence <SEQ ID 198; ORF 047.a>:

a047.pep

1 MVIIQARRGG LLVGRSIADI AQDLPDGADC QICAVYRNNR LIVPAPQTVI 51 IEGDEILFAA AAENIGAVIP ELRPKETSTR RIMIAGGGNI GYRLAKQLEH 101 AYNVKIIECR PRRAEWIAEN LDNTLVLQGS ATDETLLDNE YIDEIDVFCA

151 LTNDDESNIM SALLAKNLGA KRVIGIVNRS SYVDLLEGNK IDIVVSPHLI

201 TIGSILAHIR RGDIVAVHPI RRGTAEAIEV VAHGDKKTSA IIGRRISGIK 251 WPEGCHIAAV VRAGTGETIM GHHTETVIQD GDHIIFFVSR RRILNELEKL 301 IQVKMGFFG*

#### 96.5% identity over a 312 aa overlap m047/a047

	10	20 30	40	50	60
m047.pep	MVIIQARXXGXLVGRSIA	DIAQDLPDGADCOIC	AVYRNNRLIVPAP	OTVITEGDETLE	AΑ
a047	MVIIQARRGGLLVGRSIA				
4017	10	20 30	40		60
	10	20 30	40	30	60
	70	00	100		
0.45		80 90	100		20
m047.pep	AAENIGAVIPELRPKETO				
a047	AAENIGAVIPELRPKETS	TRRIMIAGGGNI	GYRLAKQLEHAYN	VKIIECRPRRAE	WΙ
	70	80 90	100	110	
	130 1	40 150	160	170 11	80
m047.pep	AENLDNTLVLQGSATDET				
a047	AENLDNTLVLQGSATDET				
a047	120 130				ΙV
	120 130	140 150	160	170	
		00 210	220		40
m047.pep	NRSSYVDLLEGNKIDIVV				
		111111111111111111111111111111111111111	111111111111	11111111111	1.1
a047	NRSSYVDLLEGNKIDIVV				
	180 190	200 210	220	230	
				230	
	250 2	60 270	280	290 30	00
m047.pep	TSAIIGRRISGIKWPEGO				
dedbeb	ISALIGARISGIAWELGO	HIMAVVRAGIGETIM	GHULFIAIODGDH	TIFFVSRRRILN	ىلك
0.47			1111111111111		11
a047	TSAIIGRRISGIKWPEGC			IIFFVSRRRILN	EL
	240 250	260 270	280	290	
	310				
m047.pep	EKLIQVKMGFFGX				

BNSDOCID: <WO___9957280A2_I_>

a047

EKLIQVKMGFFGX 300 310

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae

ORF 047 shows 96.2% identity over a 312 aa overlap with a predicted ORF (ORF 047.ng) from N. gonorrhoeae:

m047/g045

m047.pep	MVIIQARXXGXLVGRSIADIAQDLPDGADCQICAVYRNNRLIVPAPQTVIIEGDEILFAA	60
g047	MVIIQARRGGLLVGRSIADIAQDLPDGADCQICAVYRNNRLIVPAPQTVIIEGDEILFAA	60
m047.pep	AAENIGAVIPELRPKETQRNQPXXIMIXGGGNIGYRLAKQLEHAYNVKIIECRPRRAEWI	120
g047		117
m047.pep	AENLDNTLVLQGSATDETLLDNEYIDEIDVFCALTNDDESNIMSALLAKNLGAKRVIGIV	180
g047	AENLDNTLVLQGSATDETLLDNEYIDEIDVFCALTNDDESNIMSALLAKNLGAKRVIGIV	177
m047.pep	NRSSYVDLLEGNKIDIVVSPHLITIGSILAHIRRGDIVAVHPIRRGTAEAIEVVAHGDKK	240
g047	NRSSYVDLLEGNKIDIVVSPHLITIGSILAHIRRGDIVAVHPIRRGTAEAIEVVAHGDKK	237
m047.pep	TSAIIGRRISGIKWPEGCHIAAVVRAGTGETIMGHHTETVIQDGDHIIFFVSRRRILNEL	300
g047	TSAIIGRRISGIKWPEGCHIAAVVRAGTGETIMGHHTETVIQDGDHIIFFVSRRRILNEL	297
m047.pep	EKLIQVKMGFFGX 313	
g047	EKLIQVKMGFFGX 310	

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 199>: g048.seq

ATGCTCGACA AAGGCGAGGA GTTGCCCGTC GATTTCACCA ACCGCCTGAT 1 51 TTACTACGTC ggcCCcgTCG ATCCGGTCGG CGATGAAGTC GTCGGTCCCG 101 CAGGTCCGAC CACAGCCACC CGCATGGACA AATTTACCCG CCAAATGCTC 151 AAACAAACCG GCCTCTTGGG CATGATCGGC AAATCCGagc gcgGcgcggc 201 CACCEGCGAA GCCatCGCCG ACAACAAGGC CGTGTACCTC ATGGCAGTCG

251 GCGGCGCGC ATACCTCGTG GCAAAAGCCA TCAAATCTTC CAAAGTCTTG 301 GCGTTCCCCG AATTGGGTAT GGAAGCCGTT TACGAATTTG AAGTCAAAGA

TATGCCCGTA ACCGTCGCCG TGGACAGCAA AGGCGAATCC ATCCACGCCA CCGCCCGCG CAAATGGCAG GCGAAAATCG GCATCATCCC CGTCGAGTCT 401

451

This corresponds to the amino acid sequence <SEQ ID 200; ORF 048.ng>: g048.pep

1 MLDKGEELPV DFTNRLIYYV GPVDPVGDEV VGPAGPTTAT RMDKFTRQML 51 KQTGLLGMIG KSERGAATCE AIADNKAVYL MAVGGAAYLV AKAIKSSKVL 101 AFPELGMEAV YEFEVKDMPV TVAVDSKGES IHATAPRKWQ AKIGIIPVES

151

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 201>: m048.seq

1 ATGCTCAACA AAGGCGAAGA ATTGCCCGTC GATTTCACCA ACCGCCTGAT TTACTACGTC GGCCCCGTCG ATCCGGTCGG CGATGAAGTC GTCGGTCCGG

101	CAGGTCCGAC	CACAGCCACC	CGCATGGACA	<b>AATTCACCCG</b>	CCAAATGCTC
151	GAACAAACCG	ACCTCTTGGG	CATGATCGGC	AAATCCGAGC	GCGGCGTGGC
201	CACCTGCGAA	GCCATCGCCG	ACAACAAAGC	CGTGTACCTC	ATGGCAGTCG
251	GCGGCGCGC	GTATCTCGTG	GCAAAAGCCA	TCAAATCTTC	CAAAGTCTTG
301	GCGTTCCCCG	AATTGGGCAT	GGAAGCCATT	TACGAATTTG	AAGTCAAAGA
351	CATGCCCGTA	ACCGTCGCCG	TAGATAGCAA	AGGCGAATCC	ATCCACGCCA
	CCGCCCCGCG				
451	TGA				

### This corresponds to the amino acid sequence <SEQ ID 202; ORF 048>:

m048.pep

- MLNKGEELPV DFTNRLIYYV GPVDPVGDEV VGPAGPTTAT RMDKFTRQML 1
- 51 EQTDLLGMIG KSERGVATCE AIADNKAVYL MAVGGAAYLV AKAIKSSKVL
- 101 AFPELGMEAI YEFEVKDMPV TVAVDSKGES IHATAPRKWQ AKIGIIPVES

#### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 203>:

```
a048.seq
     1
        ATGCTCGACA AAGGCGAAGA ATTGCCCGTC GATTTCACCA ACCGCCTGAT
        TTACTACGTC GGCCCCGTCG ATCCGGTCGG CGACGAAATC GTCGGCCCAG
    51
        CAGGTCCGAC CACCGCCACC CGCATGGACA AATTCACCCG CCAAATGCTC
    101
    201 CACCTGCGAA GCCATCGCCG ACAACAAAGC CGTGTACCTC ATGGCAGTCG
    251 GCGGCGCGC GTATCTCGTG GCAAAAGCCA TCAAATCTTC CAAAGTCTTG
    301 GCGTTCCCCG AATTGGGCAT GGAAGCCATT TACGAATTTG AAGTCAAAGA
```

- 351 CATGCCCGTA ACCGTCGCCG TAGACAGCAA AGGCGAATCC ATCCACGCCA 401 CCGCCCCGCC CCAATGGCAG GCGAAAATCG GCATCATCCC CGTCAAATCT
- 451 TGA

### This corresponds to the amino acid sequence <SEQ ID 204; ORF 048.a>:

a048.pep

- 1 MLDKGEELPV DFTNRLIYYV GPVDPVGDEI VGPAGPTTAT RMDKFTRQML
- 51 EQTDLLGMIG KSERGAATCE AIADNKAVYL MAVGGAAYLV AKAIKSSKVL
- 101 AFPELGMEAI YEFEVKDMPV TVAVDSKGES IHATAPPQWQ AKIGIIPVKS

#### m048/a04896.0% identity over a 150 as overlap

	10	20	30	40	50	60
m048.pep	MLNKGEELPVDFTN			GPTTATRMDK 	FTROMLEQTE	LLGMIG
a048	MLDKGEELPVDFTN	RLIYYVGPVI		GPTTATRMDK	FTROMLEQTE	LLGMIG
	10	20	30	40	50	60
	70	80	90	100	110	120
m048.pep	KSERGVATCEAIAD	NKAVYLMAVO	GAAYLVAKAI	KSSKVLAFPE		
.040	:				111111111	
a048	KSERGAATCEAIAD	NKAVYLMAVO	GAAYLVAKAI.	KSSKVLAFPE	LGMEAIYEFE	VKDMPV
	70	80	90	100	110	120
	130	140	150			
m048.pep	TVAVDSKGESIHAT	APRKWOAKIG	SIIPVESX			
• •		11 :1[1111	1111:11			
a048	TVAVDSKGESIHAT	APPQWQAKIG	SIIPVKSX			
	130	140	150			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae

ORF 048 shows 96.4% identity over a 150 aa overlap with a predicted ORF (ORF 048.ng) from N. gonorrhoeae:

BNSDOCID: <WO 9957280A2 1 >

Santan, A salah

m048/g048 10 20 30 40 m048.pep  $\verb|MLNKGEELPVDFTNRLIYYVGPVDPVGDEVVGPAGPTTATRMDKFTRQMLEQTDLLGMIG|$ q048 MLDKGEELPVDFTNRLIYYVGPVDPVGDEVVGPAGPTTATRMDKFTRQMLKQTGLLGMIG 10 20 30 40 70 80 90 100 110 120 KSERGVATCEAIADNKAVYLMAVGGAAYLVAKAIKSSKVLAFPELGMEAIYEFEVKDMPV m048.pep KSERGAATCEAIADNKAVYLMAVGGAAYLVAKAIKSSKVLAFPELGMEAVYEFEVKDMPV g048 70 80 90 100 110 120 130 140 150 m048.pep TVAVDSKGESIHATAPRKWQAKIGIIPVESX q048 TVAVDSKGESIHATAPRKWQAKIGIIPVESX 130 140

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 205>:

ATGCGGGCGC AGGCGTTTGA TCAACCGTTC GGTCAGCTCC TGTTCGGACA

51 GGCAGAACAC TTCGCGCCGG TTGACGGCTT TCGGGTTCAG GATATTGATT

101 TGGACGGCCA TCAACGCCTC TTCCGCACCG CCTTCGCCGT TTCCGCAACC

151 CCCGTCTGCC GCCGTACCGG ATTCTGCCGC ATCGGCGTTT TCCCCGCCCT

201 CAATCTGTGC GGTTTCAAAT TCGGCACTGT CTTTTTTGGC ATCGAACCGG

251 ATTCTCCGCC GCGATTCGAT GTGTTTTTCC GAAACCGGCAC

301 AGCCTGCGG TTGAGCCAGT TTTCCTGAAG GACGATCATC GGGTCGGTTT

351 CGACTTCCTC GCCGCAATCG GCAACGGCGC EGTTGTGTC TTCCTGCCAT

This corresponds to the amino acid sequence <SEQ ID 206; ORF 049.ng>:

1 MRAQAFDQPF GQLLFGQAEH FAPVDGFRVQ DIDLDGHQRL FRTAFAVFRN 51 PVCRRTGFCR IGVFPALNLC GFKFGTVFFG IEPDSPPRFD VFFRNRHLQG 101 SLRVEPVFLK DDHRVGFDFL AAIGNGAVVF FLPFLQIRL*

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 207>:

m049.seq (partial)

1 ATGCGGGCGC AGGCGTTTGA TCAGCCGTTC GGTCAGCTCC TGTTCGGACA
51 GGCAGAACAC TTCGCGCCGG TTGACGGCTT TCGGGTTCAG GATATTGATT
101 TGGACGGCCA TCAACGTTTC TTCCGCATCG TTTTCCCCGT TTCCGAAAC
151 CGCCGGCTCA TTCGTGCCGG ATTCTGCCTC GTCGGCGTTT TCCCCGCTTT
201 CAATCTGTCC GGTTTCAAAT TCGACACTGT CTTTTTTGGT ATCAAACCGG
251 ATTCTCCGCC GCGATTCGAT GTGTTTTTCC GAAACCGACA TTTGCAGGGA
301 AGCCTGCGCG TTGAGCCAGT TTTCCTGAAG GACGATCATC GGGTCGGTTT
351 CGACTTCCTC GCCGCAATCG GCAACGGCGG CATTGTGTC CTCCTGCCAT

This corresponds to the amino acid sequence <SEQ ID 208; ORF 049>:

m049.pep (partial)

1 MRAQAFDQPF GQLLFGQAEH FAPVDGFRVQ DIDLDGHQRF FRIVFPVFRN
51 RRLIRAGFCL VGVFPAFNLS GFKFDTVFFG IKPDSPPRFD VFFRNRHLQG
101 SLRVEPVFLK DDHRVGFDFL AAIGNGGIVF LLPFFQIRL...

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 209>:

12.074

1	ATGCGGGCGC	AGGCGTTTGA	TCAGCCGTTC	GGTCAGCTCC	TGTTCGGACA
51			TTGACGGCTT		
101	TGGACGGGCA	TCAACGCTTC	TTCCGCACCG	CCTTCGCCGT	TTTCCGCAAC
151			ATTCTGCCGC		
201	CAATCTGTCC	GGTTTCAAAT	TCGGCACTGT	CTTTTTTGGC	ATCAAACCGG
251			GTGTTTTTCC		
301	AGCCTGCGCG	TTGAGCCAGT	TTTCCTGAAG	GACGATCATC	GGGTCGGTTT
351	CGACTTCCTC	GCCGCAATCG	GCAACGGCGG	CATTGTGTTC	CTCCTGCCAT
401	TTTTTCAGAT	ACGCCTT			

### This corresponds to the amino acid sequence <SEQ ID 210; ORF 049.a>:

a049.pep

1 MRAQAFDQPF GQLLFGQAEH FAPVDGFRVQ NIDLDGHQRF FRTAFAVFRN

51 PVCRRTRFCR IGVFPAFNLS GFKFGTVFFG IKPDSPPRFD VFFRNRHLQG

101 SLRVEPVFLK DDHRVGFDFL AAIGNGGIVF LLPFFQIRL

#### m049/a049 90.6% identity over a 139 aa overlap

	10	20	30	40	50	60
m049.pep	MRÄQAFDQPFGQLL	FGQAEHFAPV	/DGFRVQDIDL	DGHQRFFRIV	/FPVFRNRRL	IRAGFCL
		1111111111		111111111111111111111111111111111111111	:1 1111	1: 11
a049	MRAQAFDQPFGQLL	FGQAEHFA P\	/DGFRVQNIDL	DGHQRFFRT#	AFAVFRNPVC	RRTRFCR
	10	20	30	40	50	60
	70	80	90	100	110	120
m049.pep	VGVFPAFNLSGFKF	DTVFFGIKPI	SPPRFDVFFR	NRHLQGSLRV		
	: 1   1   1   1   1   1   1   1			1111111111		111111
a049	IGVFPAFNLSGFKF		SPPRFDVFFR	NRHLQGS <b>L</b> RV	/EPVFLKDDHI	RVGFDFL
	70	80	90	100	110	120
	130	139				
m049.pep	AAIGNGGIVFLLPF:	FQIRL				
		HHI				
a049	AAIGNGGIVFLLPF	FQIRL				
	130					

Computer analysis of this amino acid sequence gave the following results:

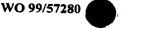
Homology with a predicted ORF from N. gonorrhoeae

ORF 049 shows 86.3% identity over a 139 aa overlap with a predicted ORF (ORF 049.ng) from N. gonorrhoeae:

m049/g049

	10	20	30	40	50	60
m049.pep	MRAQAFDQPFGQLL	<b>FGQAEHFAP</b> V	DGFRVQDIDL	DGHQRFFRIV	FPVFRNRRL	RAGECL
					1 1111	1:111
g049	MRAQAFDQPFGQLL	FGOAEHFAPV	DGFRVODIDI	DCHORLERTA		ייוון י
•	10	20	30	40	50	
	20	20	30	40	50	60
	70	0.0				
	· -	80	90	100	110	120
m049.pep	VGVFPAFNLSGFKF	DTVFFGIKPD	SPPRFDVFFR	NRHLQGSLRV	EPVFLKDDHF	RVGFDFL
					11111111111	11111
g049	IGVFPALNLCGFKF(	GTVFFGIEPD	SPPRFDVFFR	NRHLQGSLRV	EPVFLKDDH	VGFDFL
	70	80	90	100	110	120
						120
	130	139				
m049.pep	AAIGNGGIVFLLPF	FOIRL				
		:1111				
g049	AAIGNGAVVFFLPFI					
3		-				
	130	140				

BNSDOCID: <WO___9957280A2_I_>



```
The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 211>:
       g050.seq
                 atgggcgCGG GCTGGTGTCC TCCCGGCATC TTGGGCATCG GCATCGGCGG
             51
                cacgcccGAA AAAGccgtgt TGATGGcaaA AGAATCCCTG ATGAGCCACA
                TCGAcatCca aGaATTGCAG GAAAAAGCCG CGTccggggc ggaattgtcc
            101
                accaccgaAG ccCTGCGCCT cGAACTCTTT GAAAAGGTCA ACGCGCTGGG
            151
           201 CATCGGCGCG CAAGGCTTGG GCGGTCTGAC CACCGTGTTG GACGTGAAAA
           251 TCCTCGATTA CCCGACCCAT GCCGCCTCCA AACCGATTGC CATGATTCCC
           301 AACTGTGCcg ccacCCGcca cgtcgAATTT GAATTGGACG GCTCAGGtcc
           351 TGTCGAactc acgccGCcgc gtgtCGAAGA CTGA
 This corresponds to the amino acid sequence <SEQ ID 212; ORF 050.ng>:
      g050.pep
                MGAGWCPPGI LGIGIGGTPE KAVLMAKESL MSHIDIQELQ EKAASGAELS
             1
                TTEALRLELF EKVNALGIGA QGLGGLTTVL DVKILDYPTH AASKPIAMIP
            51
           101 NCAATRHVEF ELDGSGPVEL TPPRVED*
 The following partial DNA sequence was identified in N. meningitidis <SEQ ID 213>:
      m050.seq
                ATGGGCGCG GCTGGTGTCC TCCCGGCATC TTGGGTATCG GCATCGGCGG
             1
               C..agCCgAA AAAGCCGTGC TGATGGCAAA AGAGTCCCTG ATGAGCCACA
            51
           101 TCGACATTCA AGAATTGCAG GAAAAGGCCG CGTCCGGCGC GGAATTGTCC
           151 ACCACCGAAG CCCTGCGCCT CGAACTCTTT GAAAAAGTCA ACGCGCTGGG
           201 CATCGGCGCA CAAGGCTTGG GCGGACTGAC CACCGTGTTG GACGTGAAAA
           251 TCCTCGATTA TCCGACCCAC GCCGCCTCCA AACCGATTGC CATGATTCCG
           301 AACTGCGCCG CCACCCGCCA CGTCGAATTT GAATTGGACG GCTCAGGCCC
           351 TGTCGAACTC ACGCCGCCGC GCGTCGAAGA TGGCCCGATT TGA
This corresponds to the amino acid sequence <SEQ ID 214; ORF 050>:
     m050.pep
               MGAGWCPPGI LGIGIGGXAE KAVLMAKESL MSHIDIQELQ EKAASGAELS
            1
               TTEALRLELF EKVNALGIGA QGLGGLTTVL DVKILDYPTH AASKPIAMIP
            51
           101 NCAATRHVEF ELDGSGPVEL TPPRVEDGPI *
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 215>:
     a050.seg
               ATGGGCGCG GCTGGTGTCC TCCCGGCATC TTGGGCATCG GCATCGGCGG
              TACGCCCGAA AAAGCCGTGT TGATGGCGAA AGAATCCCTG ATGAGCCACA
           51
               TCGACATCCA AGAATTGCAG GAAAAAGCCG CGTCCGGCGC GGAATTGTCC
          151 ACCACCGAAG CCCTGCGCCT CGAACTCTTT GAAAAAGTCA ACGCGCTAGG
          201 CATCGGCGCG CAAGGCTTGG GCGGTCTGAC CACCGTGTTG GACGTGAAAA
          251 TCCTCGATTA CCCGACCCAC GCCGCCTCCA AACCGATTGC CATGATTCCG
          301 AACTGCGCCG CCACCCGCCA CGTCGAATTT GAATTGGACG GCTCAGGCCC
               TGTCGAACTC ACGCCGCCGC GCGTCGAAGA CTGGCCC
This corresponds to the amino acid sequence <SEQ ID 216; ORF 050.a>:
     a050.pep
               MGAGWCPPGI LGIGIGGTPE KAVLMAKESL MSHIDIQELQ EKAASGAELS
            1
               TTEALRLELF EKVNALGIGA QGLGGLTTVL DVKILDYPTH AASKPIAMIP
           51
          101 NCAATRHVEF ELDGSGPVEL TPPRVEDWP
m050/a050
            97.7% identity over a 129 aa overlap
                                    20
                                             30
                                                       40
    m050.pep
                 MGAGWCPPGILGIGIGGXAEKAVLMAKESLMSHIDIQELQEKAASGAELSTTEALRLELF
                 .
.
                 MGAGWCPPGILGIGIGGTPEKAVLMAKESLMSHIDIQELQEKAASGAELSTTEALRLELF
     a050
```

	10	20	30	40	50	60
m050.pep	70 EKVNALGIGAQGL	80 80	90 1 DVDTUAACK	100	110	120
шозо. рер	1111111111111	11111111111			111111111	33GPVEL
a050	EKVNALGIGAQGL					
	70	80	90	100	110	120
	130					
m050.pep	TPPRVEDGPIX					
a050	TPPRVEDWP					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae

ORF 050 shows 98.4% identity over a 127 aa overlap with a predicted ORF (ORF 050.ng) from N. gonorrhoeae:

m050/g050

```
10
                         20
                                 30
                                         40
                                                 50
                                                         60
          {\tt MGAGWCPPGILGIGIGGXAEKAVLMAKESLMSHIDIQELQEKAASGAELSTTEALRLELF}
m050.pep
          MGAGWCPPGILGIGIGGTPEKAVLMAKESLMSHIDIQELQEKAASGAELSTTEALRLELF
g050
                 10
                         20
                                         40
                                                 50
                                                         60
                 70
                         80
                                 90
                                        100
                                                110
                                                        120
m050.pep
          {\tt EKVNALGIGAQGLGGLTTVLDVKILDYPTHAASKPIAMIPNCAATRHVEFELDGSGPVEL}
          q050
          {\tt EKVNALGIGAQGLGGLTTVLDVKILDYPTHAASKPIAMIPNCAATRHVEFELDGSGPVEL}
                         80
                                 90
                                        100
                                                110
                                                        120
                130
m050.pep
          TPPRVEDGPIX
          9050
          TPPRVEDX
```

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 217>: g050-1.seq

1 ATGACCGTTA TCAAGCAAGA AGACTTTATT CAAAGCTATCT GCGATGCCTT

	AIGACCGIIA	ICAMGCAAGA	AGACTTTATT	CAAAGTATCT	GCGATGCCTT
51	CCAATTCATC	AGCTACTACC	ATCCAAAAGA	CTACATCGAC	GCGCTTTATA
101			AATCCCGCCG		
151			GTGTGCCGAA		
201			TCTTCCTCAA		
251	GGGATGCGGA		GAAAAGATGG		
301	GCCTACACTT		CACCCTGCGC		
351	GGCCGGCAAA	CGCCAAAACA	CCAAAGACAA	CACCCCCCCC	GTCATCCACA
401	TGAGCATCGT		AAAGTCGAAG		
451			CAAACTCGCT		
501	CATCGTCGAT		AAACCATCCC		
551			ATCGGCATCG		
601			CCTGATGAGC		TCCAAGAATT
651			GCGCGGAATT		
701			GTCAACGCGC		CGCGCAAGGC
751			GTTGGACGTG		ATTACCCGAC
801			TTGCCATGAT	TCCCAACTGT	GCCGCCACCC
851		ATTTGAATTG			
901			CGATCTGACT	TACAGCCCCG	ACAACGGCAA
951	ACGCGTCGAT				
1001	CCGGCGACGT		AACGGCAAAA		
1051	GCGCACAAAC		TATGCTCGAC		
1101			TTTACTACGT		
1151	GCGATGAAGT		GCAGGTCCGA	CCACAGCCAC	CCGCATGGAC
1201	AAATTTACCC	GCCAAATGCT	CAAACAAACC	GGCCTCTTGG	GCATGATCGG

```
1251 CAAATCCGAG CGCGGCGCGG CCACCTGCGA AGCCATCGCC GACAACAAGG
1301 CCGTGTACCT CATGGCAGTC GGCGGCGCGG CATACCTCGT GGCAAAAGCC
       1351 ATCAAATCTT CCAAAGTCTT GGCGTTCCCC GAATTGGGTA TGGAAGCCGT
      1401 TTACGAATTT GAAGTCAAAG ATATGCCCGT AACCGCCAAATGGCA GGCGAAAATC
1451 AAGGCGAATC CATCCACGCC ACCGCCCCGC GCAAATGGCA GGCGAAAATC
TTGA
 This corresponds to the amino acid sequence <SEQ ID 218; ORF 050-1.ng>:
  g050-1.pep
          1 MTVIKQEDFI QSICDAFQFI SYYHPKDYID ALYKAWQKEE NPAAKDAMTQ
       51 ILVNSRMCAE NNRPICODTG IATVFLKVGM DVOWDADMSV EKMVNEGVRR
101 AYTWEGNTLR ASVLADPAGK RONTKDNTPA VIHMSIVPGG KVEVTCAAKG
       151 GGSENKSKLA MLNPSDNIVD WVLKTIPTMG AGWCPPGILG IGIGGTPEKA
       201 VLMAKESLMS HIDIQELQEK AASGAELSTT EALRLELFEK VNALGIGAQG
       251 LGGLTTVLDV KILDYPTHAA SKPIAMIPNC AATRHVEFEL DGSGPVELTP
       301 PRVED*PDLT YSPDNGKRVD VDKLTKEEVA SWKTGDVLLL NGKILTGRDA
351 AHKRLVNMLD KGEELPVDFT NRLIYYVGPV DPVGDEVVGP AGPTTATRMD
       401 KFTRQMLKQT GLLGMIGKSE RGAATCEAIA DNKAVYLMAV GGAAYLVAKA
       451 IKSSKVLAFP ELGMEAVYEF EVKDMPVTVA VDSKGESIHA TAPRKWQAKI
       501 GIIPVES*
 g050-1/p14407
  sp|P14407|FUMB_ECOLI FUMARATE HYDRATASE CLASS I, ANAEROBIC (FUMARASE)
 >gi|280063|pir||B44511 fumarate hydratase (EC 4.2.1.2) fumB, iron-dependent - Escherichia coli
>gi|146048 (M27058) anaerobic class I fumarase (EC 4.2.1.2) [Escherichia coli] Length = 548
  Score = 172 \text{ bits } (432), \text{ Expect = } 4e-42
  Identities = 138/488 (28%), Positives = 216/488 (43%), Gaps = 22/488 (4%)
 Query: 11 QSICDAFQFISYYHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAENNRPICQDTG 70
 Q+ DA + H K L+ E + K Q L NS + A+ P CQDTG
Sbjct: 53 QAFHDASFMLRPAHQKQVAAILHDPEASEND---KYVALQFLRNSEIAAKGVLPTCQDTG 109
 Query: 71 IATVFLKVGMDVQWDADMSVEKMVNEGVRRAYTWEGNTLRASVLADPAGKRQNTKDNTPA 130
               A + K G V W E+ +++GV Y E N + A K NT N PA
 Sbjct: 110 TAIIVGKKGQRV-WTGGGD-EETLSKGVYNTYI-EDNLRYSQNAALDMYKEVNTGTNLPA 166
Query: 131 VIHMSIVPGGKVEVTCAAKGGGSENKSKL----AMLNPSDNIVDWVLKTIPTMGAGWCP 185
I + V G + + C AKGGGS NK+ L A+L P + +++++ T+G CP
 Sbjct: 167 QIDLYAVDGDEYKFLCVAKGGGSANKTYLYQETKALLTPG-KLKNFLVEKMRTLGTAACP 225
 Query: 186 PXXXXXXXXXTPEKAVLMAKESLMSHIDIQELQEKAASGAELSTTEALRLELFEKVNXXX 245
                       T + L + +H EL +
                                                          +
                                                                 L EL E+
Sbjct: 226 PYHIAFVIGGTSAETNLKTVKLASAHY-YDELPTEGNEHGQAFRDVQLEQELLEEAQKLG 284
Query: 246 XXXXXXXXXTTVLDVKILDYPTHAASKPIAMIPNCAATRHVEFELDGSG----PVELTPP 301
                            D++++ P H AS P+ M +C+A R+++ +++ G +E P
Sbjct: 285 LGAQFGGKYFAH-DIRVIRLPRHGASCPVGMGVSCSADRNIKAKINREGIWIEKLEHNPG 343
uery: 302 RVEDXPDLTYSPDNGKRVDVDKLTKE---EVASWKTGDVLLLNGKILTGRDAAHKRLVNM 358
+ +VD+++ KE +++ + L L G I+ GRD AH +L +
                                                       L L G I+ GRD AH +L +
Sbjct: 344 QYIPQELRQAGEGEAVKVDLNRPMKEILAQLSQYPVSTRLSLTGTIIVGRDIAHAKLKEL 403
Query: 359 LDKGEELPVDFTNRLIYYXXXXXXXXXXXXXXXXTTATRMDKFTRQMLKQTGLLGMIGK 418
             +D G+ELP + IYY
                                                     TTA RMD + + G + M+ K
Sbjct: 404 IDAGKELPQYIKDHPIYYAGPAKTPAGYPSGSLGPTTAGRMDSYVDLLQSHGGSMIMLAK 463
Query: 419 SERGAATCEAIADNKAVYLMAVGG-AAYLVAKAIKSSKVLAFPELGMEAVYEFEVKDMPV 477
                     +A + YL ++GG AA L ++IK + +A+PELGMEA+++ EV+D P
Sbjct: 464 GNRSQQVTDACHKHGGFYLGSIGGPAAVLAQQSIKHLECVAYPELGMEAIWKIEVEDFPA 523
Query: 478 TVAVDSKG 485
              + VD KG
Sbjct: 524 FILVDDKG 531
```

#### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 219>: m050-1.seq

- 1 ATGACCGTCA TCAAACAGGA AGACTTTATC CAAAGCATTT GCGATGCCTT 51 CCAATTCATC AGCTACTATC ATCCCAAAGA CTACATCGAC GCGCTTTATA
- 101 AGGCGTGGCA GAAGGAAGAA AATCCTGCCG CCAAAGACGC GATGACGCAG
- ATTTTGGTCA ACAGCCGTAT GTGTGCGGAA AACAACCGCC CCATCTGCCA
  AGACACAGGT ATCGCAACCG TCTTCCTCAA AGTCGGTATG AACGTCCAAT
  GGGATGCGGA CATGAGCGTG GAAGAGATGG TTAACGAAGG CGTACGCCGC

301	GCCTACACTT	GGGAAGGCAA	TACGCTGCGC	GCTTCCGTCC	TCGCCGATCC
351	GGCCGGCAAA	CGCCAAAACA	CCAAAGACAA	CACCCCCGCC	GTCATCCATA
401	TGAGCATCGT	GCCGGGCGGT	AAAGTCGAAG	TAACCTGCGC	GGCAAAAGGC
451	GGCGGCTCTG	AAAACAAATC	CAAACTCGCC	ATGCTCAATC	CTTCCGACAA
501	CATCGTCGAT	TGGGTATTGA	AAACCATCCC	GACCATGGGC	GCGGGCTGGT
551	GTCCTCCCGG	CATCTTGGGT	ATCGGCATCG	GCGGCACGCC	CGAAAAAGCC
601	GTGCTGATGG	CAAAAGAGTC	CCTGATGAGC	CACATCGACA	TTCAAGAATT
651	GCAGGAAAAG	GCCGCGTCCG	GCGCGGAATT	GTCCACCACC	GAAGCCCTGC
701	GCCTCGAACT	CTTTGAAAAA	GTCAACGCGC	TGGGCATCGG	CGCACAAGGC
751	TTGGGCGGAC	TGACCACCGT	GTTGGACGTG	AAAATCCTCG	ATTATCCGAC
801	CCACGCCGCC	TCCAAACCGA	TTGCCATGAT	TCCGAACTGC	GCCGCCACCC
851	GCCACGTCGA	ATTTGAATTG	GACGGCTCAG	GCCCTGTCGA	ACTCACGCCG
901	CCGCGCGTCG	AAGACTGGCC	CGATTTGACT	TACAGCCCCG	ACAACGGCAA
951	ACGCGTCGAT	GTCGACAAGC	TGACCAAAGA	AGAAGTGGCA	AGCTGGAAAA
1001	CCGGCGACGT	ATTGCTGTTG	AACGGCAAAA	TCCTCACCGG	CCGCGATGCC
1051	GCACACAAAC	GCCTCGTCGA	TATGCTCAAC	AAAGGCGAAG	AATTGCCCGT
1101	CGATTTCACC	AACCGCCTGA	TTTACTACGT	CGGCCCCGTC	GATCCGGTCG
1151	GCGATGAAGT	CGTCGGTCCG	GCAGGTCCGA	CCACAGCCAC	CCGCATGGAC
1201	AAATTCACCC	GCCAAATGCT	CGAACAAACC	GACCTCTTGG	GCATGATCGG
1251	CAAATCCGAG	CGCGGCGTGG	CCACCTGCGA	AGCCATCGCC	GACAACAAAG
1301	CCGTGTACCT	CATGGCAGTC	GGCGGCGCGG	CGTATCTCGT	GGCAAAAGCC
1351	ATCAAATCTT	CCAAAGTCTT	GGCGTTCCCC	GAATTGGGCA	TGGAAGCCAT
1401	TTACGAATTT	GAAGTCAAAG	ACATGCCCGT	AACCGTCGCC	GTAGATAGCA
1451	AAGGCGAATC	CATCCACGCC	ACCGCCCCCC	GCAAATGGCA	GGCGAAAATC
1501	GGCATCATCC	CCGTCGAATC	TTGA		

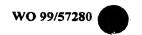
# This corresponds to the amino acid sequence <SEQ ID 220; ORF 050-1>: m050-1.pep

1	MTVIKQEDFI	QSICDAFQFI	SYYHPKDYID	ALYKAWQKEE	NPAAKDAMTQ
51	ILVNSRMCAE	NNRPICQDTG	IATVFLKVGM	NVQWDADMSV	EEMVNEGVRR
101	AYTWEGNTLR	ASVLADPAGK	RONTKONTPA	VIHMSIVPGG	KVEVTCAAKG
151	GGSENKSKLA	MLNPSDNIVD	WVLKTIPTMG	AGWCPPGILG	IGIGGTPEKA
201	VLMAKESLMS	HIDIQELQEK	AASGAELSTT	EALRLELFEK	VNALGIGAQG
251	LGGLTTVLDV	KILDYPTHAA	SKPIAMIPNC	AATRHVEFEL	DGSGPVELTP
301	PRVEDWPDLT	YSPDNGKRVD	VDKLTKEEVA	SWKTGDVLLL	NGKILTGRDA
351	AHKRLVDMLN	KGEELPVDFT	NRLIYYVGPV	DPVGDEVVGP	<b>AGPTTATRMD</b>
401	KFTRQMLEQT	DLLGMIGKSE	RGVATCEAIA	DNKAVYLMAV	GGAAYLVAKA
451	IKSSKVLAFP	ELGMEAIYEF	EVKDMPVTVA	VDSKGESIHA	TAPRKWQAKI
501	GIIPVES*				

#### m050-1/g050-1 98.2% identity in 507 aa overlap

m050-1.pep g050-1	10 20 MTVIKQEDFIQSICDAFQFI:		ENPAAKDAMTQILVNSRMCAE
m050-1.pep g050-1	70 80 NNRPICQDTGIATVFLKVGMI	:11111111111111111111111111111111111111	RAYTWEGNTLRASVLADPAGK               RAYTWEGNTLRASVLADPAGK
m050-1.pep g050-1	130 140 RQNTKDNTPAVIHMSIVPGGI IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	111111111111111111111111111111111111111	MLNPSDNIVDWVLKTIPTMG
m050-1.pep g050-1	190 200 AGWCPPGILGIGIGGTPEKAV                      AGWCPPGILGIGIGGTPEKAV 190 200		KAASGAELSTTEALRLELFEK
m050-1.pep g050-1	250 260 VNALGIGAQGLGGLTTVLDVI	1111111111111111111111	CAATRHVEFELDGSGPVELTP

BNSDOCID: <WO___9957280A2_i_>



m050-1.pep g050-1	310 PRVEDWPDLTYSP	1111111111	111111111	11111111111	111111111	111:11:
m050-1.pep	370 KGEELPVDFTNRL !!!!!!!!! KGEELPVDFTNRL: 370		111111111.		111:11 11	111111
m050-1.pep g050-1	430 RGVATCEAIADNKA   :         RGAATCEAIADNKA 430				11:11:11	HILLIA
m050-1.pep	490 VDSKGESIHATAPF            VDSKGESIHATAPF 490	1111111111	111			

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 221>: a050-1.seq

```
1 ATGACCGTCA TCAAACAGGA AGACTTTATC CAAAGCATTT GCGATGCCTT
   51 CCAATTCATC AGCTACTACC ATCCCAAAGA CTACATCGAC GCGCTTTATA
 101 AGGCGTGGCA GAAGGAAGAA AACCCCGCCG CCAAAGACGC GATGACGCAG
151 ATTTTGGTCA ACAGCCGCAT GTGTGCCGAA AACAACCGCC CCATCTGCCA
  201 AGATACCGGT ATCGCGACCG TGTTTTTGAA AGTCGGTATG GATGTGCAAT
       GGGATGCAGA CATGAGCGTC GAAGAGATGG TTAACGAAGG CGTGCGCCGC
  251
  301 GCCTACACTT GGGAAGGCAA TACGCTGCGC GCTTCCGTTC TCGCCGACCC
  351
       CGCCGGCAAA CGCCAAAATA CCAAAGACAA CACGCCCGCC GTCATCCATA
 401 TGAGCATCGT GCCGGGCGAC AAAGTCGAAG TAACCTGCGC GGCAAAAGGC
451 GGCGGTTCTG AAAACAAATC CAAACTCGCC ATGCTCAACC CTTCCGACAA
 501 CATCGTCGAT TGGGTATTGA AAACCATTCC GACCATGGGC GCGGGCTGGT
551 GTCCTCCCGG CATCTTGGGC ATCGGCATCG GCGGTACGCC CGAAAAAGCC
 601 GTGTTGATGG CGAAAGAATC CCTGATGAGC CACATCGACA TCCAAGAATT
651 GCAGGAAAAA GCCGCGTCCG GCGCGGAATT GTCCACCACC GAAGCCCTGC
 701 GCCTCGAACT CTTTGAAAAA GTCAACGCGC TAGGCATCGG CGCGCAAGGC
 751 TTGGGCGGTC TGACCACCGT GTTGGACGTG AAAATCCTCG ATTACCCGAC
801 CCACGCCGCC TCCAAACCGA TTGCCATGAT TCCGAACTGC GCCGCCACCC
       GCCACGTCGA ATTTGAATTG GACGGCTCAG GCCCTGTCGA ACTCACGCCG
 851
       CCGCGCGTCG AAGACTGGCC CGATTTGACT TACAGCCCCG ACAACGGCAA
 901
 951
       ACGCGTCGAT GTCGACAAGC TGACCAAAGA AGAAGTGGCA AGCTGGAAAA
1001
       CCGGCGACGT ATTGCTGTTG AACGGCAAAA TCCTCACCGG CCGCGATGCC
1051 GCACACAAAC GCCTCGTCGA TATGCTCGAC AAAGGCGAAG AATTGCCCGT
1101
       CGATTTCACC AACCGCCTGA TTTACTACGT CGGCCCCGTC GATCCGGTCG
1151 GCGACGAAAT CGTCGGCCCA GCAGGTCCGA CCACCGCCAC CCGCATGGAC
1201 AAATTCACCC GCCAAATGCT CGAACAAACC GACCTCTTGG GCATGATCGG
1251 CAAATCCGAG CGCGGCGCGG CCACCTGCGA AGCCATCGCC GACAACAAG
1301 CCGTGTACCT CATGGCAGTC GGCGGCGCGG CGTATCTCGT GGCAAAAGCC
1351 ATCAAATCTT CCAAAGTCTT GGCGTTCCCC GAATTGGGCA TGGAAGCCAT
      TTACGAATTT GAAGTCAAAG ACATGCCCGT AACCGTCGCC GTAGACAGCA
1401
1451
      AAGGCGAATC CATCCACGCC ACCGCCCCGC CCCAATGGCA GGCGAAAATC
1501 GGCATCATCC CCGTCAAATC TTGA
```

### This corresponds to the amino acid sequence <SEQ ID 222; ORF 050-1.a>: a050-1.pep

1	MIVIKQEDFI	QSICDAFQFI	SYYHPKDYID	ALYKAWOKEE	NPAAKDAMTO
51	ILVNSRMCAE	NNRPICQDTG	IATVFLKVGM	DVOWDADMSV	FEMUNECUED
101	AYTWEGNTLR	ASVLADPAGK	RQNTKDNTPA	VIHMSTUDGO	RABAMEGAVV
151	GGSENKSKLA	MLNPSDNIVD	WVLKTIPTMG	AGWCDDGTIC	TOTOCHDENA
201	VLMAKESLMS	HIDIOELOEK	AASGAELSTT	EVIDIE! EEE	TGIGGIPEKA
251	LGGLTTVLDV	KILDYPTHAA	SKPIAMIPNC	PUPLIFIED	VNALGIGAQG
301	PRVEDWPDLT	AZDDNCKDAD	VDKLTKEEVA	CHRECONTE	DGSGPVELTP
351	AHKRLVDMLD	RCEEI DUDEM	VDALIALEVA	SWATGDVLLL	NGKILTGRDA
401	KETROMI FOT	DITCMICACE	NKLITIVGPV	DPVGDEIVGP	AGPTTATRMD
451	KFTRQMLEQT	DELCHIER THER	RGAATCEAIA	DNKAVYLMAV	GGAAYLVAKA
501	IKSSKVLAFP GIIPVKS*	ELGMEAIIEF	EVKDMPVTVA	VDSKGESIHA	TAPPQWQAKI

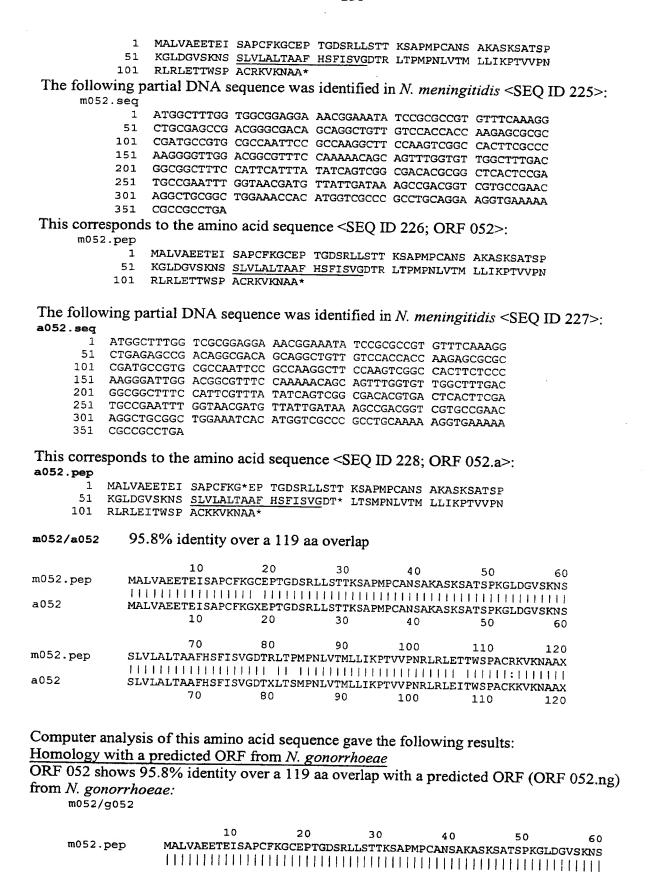
***

```
98.4% identity in 507 aa overlap
a050-1/m050-1
                                30
                                        40
                                                50
         MTVIKQEDFIQSICDAFQFISYYHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAE
a050-1.pep
          m050-1
          MTVIKQEDFIQSICDAFQFISYYHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAE
                10
                        20
                                30
                                        40
                                                50
                70
                        80
                                90
                                        100
                                               110
a050-1.pep
         NNRPICQDTGIATVFLKVGMDVQWDADMSVEEMVNEGVRRAYTWEGNTLRASVLADPAGK
          NNRPICQDTGIATVFLKVGMNVQWDADMSVEEMVNEGVRRAYTWEGNTLRASVLADPAGK
m050-1
                        80
                                90
                                        100
                                               110
               130
                       140
                               150
                                        160
                                               170
                                                       180
         RONTKONTPAVIHMSIVPGDKVEVTCAAKGGGSENKSKLAMLNPSDNIVDWVLKTIPTMG
a050-1.pep
          m050-1
          RQNTKDNTPAVIHMSIVPGGKVEVTCAAKGGGSENKSKLAMLNPSDNIVDWVLKTIPTMG
               130
                       140
                                       160
                                               170
                                                       180
               190
                       200
                               210
                                       220
                                               230
         AGWCPPGILGIGIGGTPEKAVLMAKESLMSHIDIQELQEKAASGAELSTTEALRLELFEK
a050-1.pep
          m050-1
          AGWCPPGILGIGGGTPEKAVLMAKESLMSHIDIQELQEKAASGAELSTTEALRLELFEK
                       200
                               210
               190
                                       220
                                               230
               250
                       260
                               270
                                       280
                                               290
a050-1.pep
         VNALGIGAOGLGGLTTVLDVKILDYPTHAASKPIAMIPNCAATRHVEFELDGSGPVELTP
          m050 - 1
          VNALGIGAQGLGGLTTVLDVKILDYPTHAASKPIAMIPNCAATRHVEFELDGSGPVELTP
               250
                       260
                               270
                                       280
                                               290
                       320
               310
                               330
                                       340
                                               350
a050-1.pep
         PRVEDWPDLTYSPDNGKRVDVDKLTKEEVASWKTGDVLLLNGKILTGRDAAHKRLVDMLD
          m050-1
          PRVEDWPDLTYSPDNGKRVDVDKLTKEEVASWKTGDVLLLNGKILTGRDAAHKRLVDMLN
               310
                       320
                               330
                                       340
                                               350
                                                       360
               370
                       380
                               390
                                        400
                                               410
                                                       420
         KGEELPVDFTNRLIYYVGPVDPVGDEIVGPAGPTTATRMDKFTRQMLEQTDLLGMIGKSE
a050-1.pep
          m050-1
         KGEELPVDFTNRLIYYVGPVDPVGDEVVGPAGPTTATRMDKFTRQMLEQTDLLGMIGKSE
               370
                       380
                               390
                                        400
                                                       420
               430
                       440
                               450
                                       460
                                               470
         {\tt RGAATCEAIADNKAVYLMAVGGAAYLVAKAIKSSKVLAFPELGMEAIYEFEVKDMPVTVA}
a050-1.pep
          m050-1
          RGVATCEAI ADNKAVYLMAVGGAAYLVAKA I KSSKVLAFPELGMEA I YEFEVKDMPVTVA
               430
                       440
                               450
                                       460
                                               470
                                                       480
               490
                       500
a050-1.pep
         VDSKGESIHATAPPQWQAKIGIIPVKSX
          111111111111111111111111111111111111
m050-1
         VDSKGESIHATAPRKWQAKIGIIPVESX
               490
                       500
```

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 223>:

g052.seq ATGGCTTTGG TGGCGGAGGA AACGGAAATA TCCGCGCCGT GTTTCAAAGG CTGCGAGCCG ACGGGCGACA GCAGGCTGTT GTCCACCACC AAGAGCGCGC CGATGCCGTG CGCCAATTCC GCCAAGGCTT CCAAGTCGGC CACTTCGCCC AAGGGGTTGG ACGGCGTTTC CAAAAACAGC AGTTTGGTGT TGGCTTTGAC GGCGGCTTTC CATTCATTTA TATCAGTCGG CGACACGCGG CTCACTCCGA TGCCGAATTT GGTAACGATG TTATTGATAA AGCCGACGGT CGTGCCGAAC AGGCTGCGGC TGGAAACCAC ATGGTCGCCC GCCTGCAGGA AGGTGAAAAA CGCCGCCTGA 

This corresponds to the amino acid sequence <SEQ ID 224; ORF 052.ng>: g052.pep



g052	MALVAEETEISAPC	KGCEPTGDS	RLLSTTKSAP	MPCANSAKAS	KSATSPKGLI	GVSKNS
	10	20	30	40	50	60
	70	80	90	100	110	120
m052.pep	SLVLALTAAFHSFIS	SVGDTWLTSM	PNLATMLLIK	PTVVPNRLRI	EITWSPACK	TVKNAAX
			111:111111	1111111111	1 111111:1	HHHH
g052	SLVLALTAAFHSFIS	SVGDTRLTPM	PNLVTMLLIK	PTVVPNRLRI	ETTWSPACRE	VKNAAX
	70	80	90	100	110	120

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 229>: g073.seq

```
1 ATGTGTATGC CATACGCAAT AAGGGTTTCA GACGGCATCT GCCGCATTTT
```

- 51 TCCGCCGATG CCGTCTGAAA CACGCAATCA GCGCGCGAGT GCCTGTTTCA
- 101 AATCGTCAAT CAAATCGCCA ACATATTCCA AACCGACCGA CAGGCGCACC
- 151 AGTCCGGGGC GGatacCGGC GGCGAGTTTT TCTTCGGGCT GCATCCTGCC
- 201 GTGCGTGGTT GTCCACGGAT TGGTGATGGT CGAGCGCACG TCGCCGAGGT
  251 TGGCGGTACG GGAAAAGAGT TCCACGACTT TCCACGCGGC TGCTTGGTCG
- 301 GCGACTTCAA AACCGATGAC GATGCCGCCG CCGTTTTGCT GTTTGCGGAT
- 351 AAGCTCCGCC TGCGGATGGT CGGGCAATCC GGTGTAG

This corresponds to the amino acid sequence <SEQ ID 230; ORF 073.ng>: g073.pep

- 1 MCMPYAIRVS DGICRIFPPM PSETRNORAS ACFKSSIKSP TYSKPTDRRT
- 51 SPGRIPAASF SSGCILPCVV VHGLVMVERT SPRLAVREKS STTFHAAAWS.
- 101 ATSKPMTMPP PFCCLRISSA CGWSGNPV*

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 231>:

m073.seq

- 1 ATGTGTATGC CATATAAGAT AAGGGTTTCA GACGGCATCT GCTGTCCAAT
- 51 GCCGTCTGAA ACACGCAATC AGCGTGCGAG TGCCTGTTTC AAATCGTCAA
- 101 TCAAATCGCC AACATATTCC AAACCGACCG ACAGGCGCAC CAATCCGGGG
- 151 CGGATGTTGG CGGCGAGTTT TTCTTCGGGC TGCATCCTGC CGTGCGTGGT
- 201 TGTCCACGGG TGGGTAATGG TCGAGCGCAC GTCACCGAGG TTGGCGGTGC
- 251 GGGAAAAGAG TTCCACGCCG TCCACAACTT TCCACGCCGC TTCTTGATCG
- 301 GCAACTTCAA AGCCGATGAC GATGCCGCCG CCGTTTTGCT GTTTGCGGAT
- 351 AAGCGCCGCC TGAGGATGGT CGGACAATCC GGTGTAG

This corresponds to the amino acid sequence <SEQ ID 232; ORF 073>:

m073.pep

- 1 MCMPYKIRVS DGICCPMPSE TRNQRASACF KSSIKSPTYS KPTDRRTNPG
- 51 RMLAASFSSG CILPCVVVHG WVMVERTSPR LAVREKSSTP STTFHAASXS
- 101 ATSKPMTMPP PFCCLRISAA XGWSDNPV*

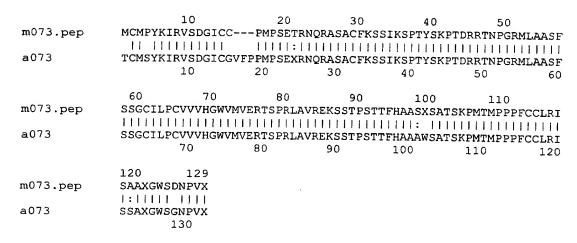
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 233>: a073.seq

- 1 ACGTGTATGT CATATAAGAT AAGGGTTTCA GACGGCATTT GCGGTGTTTT
- 51 TCCGCCGATG CCGTCTGAA. CACGCAATCA GCGCGCGAGT GCCTGTTTCA
- 101 AATCGTCAAT CAAATCGCCA ACATATTCCA AACCGACCGA CAGGCGCACC
- 151 AATCCGGGGC GGATGTTGGC GGCGAGTTTT TCTTCGGGCT GCATCCTGCC
- 201 GTGCGTGGTT GTCCACGGAT GGGTAATGGT CGAGCGCACG TCGCCGAGGT
- 251 TGGCGGTACG GGAGAAAAGT TCGACGCCGT CCACGACTTT CCACGCGGCT 301 GCTTGGTCGG CGACTTCAAA GCCGATGACG ATGCCGCCGC CGTTTTGCTG
- 351 TTTGCGGATA AGCTCCGCCT GAGGATGGTC GGGTAATCCG GTGTAA

This corresponds to the amino acid sequence <SEQ ID 234; ORF 073.a>: a073.pep

- 1 TCMSYKIRVS DGICGVFPPM PSEXRNQRAS ACFKSSIKSP TYSKPTDRRT
- 51 NPGRMLAASF SSGCILPCVV VHGWVMVERT SPRLAVREKS STPSTTFHAA
- 101 AWSATSKPMT MPPPFCCLRI SSA*GWSGNP V*

m073/a073 92.3% identity over a 130 aa overlap



Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae

ORF 073 shows 87.0% identity over a 131 aa overlap with a predicted ORF (ORF 073.ng) from N. gonorrhoeae:

m073/g073

	10	20	30	40	50
m073.pep	MCMPYKIRVSDGICC	PMPSETRN	QRASACFKS	SIKSPTYSKPT	DRRTNPGRMLAASF
		111111			1111:111: 1111
g073	MCMPYAIRVSDGICR	IFPPMPSETRN	<b>QRASACFKS</b>	SIKSPTYSKPT	DRRTSPGRIPAASF
	10	20	30	40	50 60
	60 70	0.0			
	· -	80	90	100	110
m073.pep	SSGCILPCVVVHGWV	MVERTSPRLAV	REKSSTPST	TFHAASXSATS	KPMTMPPPFCCLRI
		1111111111		111111111111111111111111111111111111111	
g073	SSGCILPCVVVHGLV	MVEDTG DDT.AV	DEKCCT	TELLA A MOAMO:	111111111111
3					KPMTMPPPFCCLRI
	70	80	90	100	110
	120 129				
m072 man					
m073.pep	SAAXGWSDNPVX				
	1:				
g073	SSACGWSGNPVX				
	120				

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 235>: g075.seq

- 1 ATGCCGCCTT ACTTCATCAC CCTCTTAACG ATGGAAAATA CAAAAAGCGC
  51 GGCGAAAACG CCCACTACAA TCCAACCGGC TTCCATACCG TCCGCTTTTG
  101 CGGCTTCCAA AGCGTTTTTT GCCGTTTCGG GCAACGCTGC GTTTGCCTGT
  151 GCCGCCAAAG CCAGCGGGGC GGCTGTTACA ACAGCCAGTT TTGCGCCGTA
  201 TTTACGGCAG GTGTTAATAA ATTTCATGAT ATTTTCCTTT ACGAAATTTT
  251 TAAAAAAATG TGTTTGCGGG CTTTGTGAAG GTTTTAGAGA CCGCCTGCCG
  301 GGCCTCTTAA ACTTAATCTT CTTTTTCGTA GAATCCGAAA ATTACAAATT
  351 CCCCGCCTAT CTCTTCCAAT GCCGAGCTAA AAGCGTCTTC ATAGCTGTCA
- This corresponds to the amino acid sequence <SEQ ID 236; ORF 075.ng>: g075.pep
  - 1 MPPYFITLLT MENTKSAAKT PTTIQPASIP SAFAASKAFF AVSGNAAFAC 51 AAKASGAAVT TASFAPYLRQ VLINFMIFSF TKFLKKCVCG LCEGFRDRLP
  - 101 GLLNLIFFFV ESENYKFPAY LFQCRAKSVF IAVIFTG*

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 237>: m075.seg

- ATGCCGTCTT ACTTCATCAC TCTCTTAACG ATGGAAAATA CAAAAAGCGC
- 51 GGCGAAAATG CCCACTACAA TCCAACCGGC TTCCATACCG TCCGCTTTTG
- 101 CGGCTTCCAA AGCGTTTTTT GCCGTATCGG GCAACGTTGC ATTTGCATGT
- 151 GCGGCCAAAG CCAGGGGAGC AGCTGTTACA ACAGCCAGTT TTGCGCCGTA
- 201 TTTACGGCAG GTGTTAATAA ATTTCATGAT ATTTTCCTTC AAAAAGTGTT
- 251 TGGCGGTAAT GGATGGAGCG TTTTTCAGAC GACCGCCGAA CATCCGAAAA
- 301 TCAGTCTTTC AAAAATCCGA ATACGACAAA TTCGTATTGG TTGCCGATTT
- 351 CTTCCAAACC TO 401 TTGGTGATTA A CTTCCAAACC TGCGTTAATC GCTTCTTCGA AGTCGTAGAA ATAATCGGCA

This corresponds to the amino acid sequence <SEQ ID 238; ORF 075>:

- m075.pep
  - MPSYFITLLT MENTKSAAKM PTTIQPASIP SAFAASKAFF AVSGNVAFAC 1
  - 51 AAKARGAAVT TASFAPYLRQ VLINFMIFSF KKCLAVMDGA FFRRPPNIRK
  - 101 SVFQKSEYDK FVLVADFFQT CVNRFFEVVE IIGIGD*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae

ORF 075 shows 65.7% identity over a 137 aa overlap with a predicted ORF (ORF 075.ng) from N. gonorrhoeae:

m075/g075

	10	20	30	40	50	60
m075.pep	MPSYFITLLTMENT	KSAAKMPTT	IQPASIPSAFAA	SKAFFAVSG	<b>WAFACAAKA</b>	RGAAVT
	11 1111111111				:	11111
g075	MPPYFITLLTMENT	KSAAKTPTT	IOPASIPSAFAA	SKAFFAVSG		SGAAVT
<b>3</b>	10	20	30	40	50	60
	10	20	50	40	30	80
	70	80	90	100	110	1
m075.pep	TASFAPYLRQVLIN	MIFSF	-KKCLAVMDGAF	FRRPPNIRKS	VFOKSEYDK	TVLVAD
			111: : :1			1
q075	TASFAPYLRQVLIN	MTESETKE			, ,	VEDAV
9075	· ·					
	70	80	90	100	110	120
	100 100					. ,
	120 130					
m075.pep	FFQTCVNRFFEVVE:	IIGIGDX				
	- :     - :   :   :	:				
q075	LFOCRAKSVFIAVII	TGX				
_	130					

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 239>: a075.seq

- ATGCCGTCTT ACTTCATCAC TCTCTTAACG ATGGAAAAGA CAAAAAGCGC 51 GGCGAAAACG CCCACTACAA TCCAACCGGC TTCCATACCG TCCGCTTTTG CGGCTTCCAA AGCGTTTTTT GCTGTATCGG GCAACGTTGC ATTTGCATGT 101 GCGGCCAAAG CCAGGGGAGC AGCTGTTACA ACAGCCAGTT TTGCGCCGTA
- 201 TTTACGGCAG GTGTTAATAA ATTTCATGAT ATTTTCCTTC AAAAAGTGTT
- 251 TGGCGGTAAT GGATGGAGCG TTTTTCAGAC GACCGCCGAA CATCCGAAAA
- 301 TCAGTCTTC AAAAATCCGA ATACGACAAA TTCGTATTGG TTGCCGATTT
- CTTCCAAACC TGCGTTAATC GCTTCTTCGA AGTCGTAGAA ATAATCGGCA 351
- 401 TTGGTGATTA A

This corresponds to the amino acid sequence <SEQ ID 240; ORF 075.a>: a075.pep

- 1 MPSYFITLLT MEKTKSAAKT PTTIQPASIP SAFAASKAFF AVSGNVAFAC
- 51 AAKARGAAVT TASFAPYLRQ VLINFMIFSF KKCLAVMDGA FFRRPPNIRK
- SVFQKSEYDK FVLVADFFQT CVNRFFEVVE IIGIGD*



WO 99/57280

```
m075/a075
              98.5% identity over a 136 aa overlap
                                                  40
             MPSYFITLLTMENTKSAAKMPTTIQPASIPSAFAASKAFFAVSGNVAFACAAKARGAAVT
 m075.pep
              a075
              MPSYFITLLTMEKTKSAAKTPTTIQPASIPSAFAASKAFFAVSGNVAFACAAKARGAAVT
                                         30
                                                  40
                                                            50
                     70
                               80
                                         90
                                                 100
                                                           110
             TASFAPYLRQVLINFMIFSFKKCLAVMDGAFFRRPPNIRKSVFQKSEYDKFVLVADFFQT
 m075.pep
              TASFAPYLRQVLINFMIFSFKKCLAVMDGAFFRRPPNIRKSVFQKSEYDKFVLVADFFQT
 a075
                     70
                               80
                                        90
                                                 100
                    130
m075.pep
             CVNRFFEVVEIIGIGDX
             2075
             CVNRFFEVVEIIGIGDX
                    130
The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 241>:
     q080.seq
               ATGTGGGATA ATGCCGAAGC GATGGAACGG CTGACGCGCT GGCTGCTTGT
            1
               CATGATGGCG ATGCTGCTTG CTGCGTCCGG GCTGGTTTGG TTTTACAATT
           51
          101 CGAATCATCT GCCCGTCAAG CAGGTGTCGC TGAAGGGCAA CCTGGTTTAT
          151 TCCGATAAGA AGGCATTGGG CAGTTTGGCG AAAGAATACA TCCATGGGAA
          201 TATTTTGAGG ACGGACATCA ATGGCGCACA GGAAGCCTAC CGCCGGTATC
          251 CGTGGATTGC GTCGGTCATG GTGCGCCGCC GTTTTCCCGA TACGGTTGAG
          301 GTCGTCCTGA CCGAGCGCAA GCCGGTTGCA CGTTGGGGCG ACCATGCCTT
          351 GGTGGACGGC GAAGGCAATG TTTTTGAAGC CCGCTTGGAC AGACCCGGAA
          401 TGCCGGTATT CAGAGGCGCG GAAGGAACGT CTGCCGAAAT GCTCCGCCGT
          451 TATGACGAAT TTTCGACTGT TTTGGCAAAA CAGGGTTTGG GCATCAAAGA
               GATGACCTAT ACGGCACGTT CGGCGTGGAA TGTCGTTTTG GACAACGGCA
               TCACCGTCAG GCTCGGACGG GAAAACGAGA TGAAACGCCT CCGGCTTTTT
              ACCGAAGCGT GGCAGCATCT GttgcGTAAG AATAAAAATC GGTTATCCTA
          601
              TGTGGATATG Aggtataagg acggatttTC agtcccccat gctCCCGACG
          651
          701 GTTTACCCGA AAAAGAATCC GAAGAATatt gggaacaggt ttgggacata
          751 ttacggcctg gcgtcggaaa cggttcgacg caaatttcaa tcagttatAA
              GGGCAGacga acaatggaac AGcagtaa
This corresponds to the amino acid sequence <SEQ ID 242; ORF 080.ng>:
     g080.pep
              MWDNAEAMER LTRWLLVMMA MLLAASGLVW FYNSNHLPVK QVSLKGNLVY
              SDKKALGSLA KEYIHGNILR TDINGAQEAY RRYPWIASVM VRRRFPDTVE
           51
              VVLTERKPVA RWGDHALVDG EGNVFEARLD RPGMPVFRGA EGTSAEMLRR
          101
         151 YDEFSTVLAK QGLGIKEMTY TARSAWNVVL DNGITVRLGR ENEMKRLRLF
              TEAWQHLLRK NKNRLSYVDM RYKDGFSVPH APDGLPEKES EEYWEQVWDI
         251 LRPGVGNGST QISISYKGRR TMEQQ*
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 243>:
     m080.seq
              ATGTGGGATA ATGCCGAAGC GATGGAACGG CTGACGCGCT GGCTGCTTGT
           1
              CATGATGGCG ATGCTGCTTG CTGCGTCCGG GCTGGTTTGG TTTTACAATT
          51
              CGAATCATCT GCCCGTCAAG CAGGTGTCGC TGAAGGGCAA CCTGGTTTAT
         101
              TCCGATAAGA AGACATTGGG CAGTTTGGCG AAAGAATACA TCCATGGGAA
         151
              TATTTTGAGG ACGGACATCA ATGGCGCACA GGAGGCCTAC CGCCGGTATC
         201
              CGTGGATTGC GTCGGTCATG GTGCGCCGCC GTTTTCCCGA CACGGTTGAG
         251
              GTCGTCCTGA CCGAGCGCAA GCCGGTCGCG CGTTGGGGCG ACCATGCCTT
         301
              GGTGGACGGC GAAGGCAATG TTTTTGAAGC CCGCTTGGAC AGACCCGGAA
         351
              TGCCGGTATT CAGAGGCGCG GAAGGAACGT CTGCCGAAAT GCTCCGCCGT
         401
              TATGACGAAT TTTCGACTGT TTTGGCAAAA CAGGGTTTGG GCATCAAAGA
         451
         501 GATGACCTAT ACGGCACGTT CGGCGTGGAT TGTCGTTTTG GACAACGGCA
         551 TCACCGTCAG GCTCGGACGG GAAAACGAGA TGAAACGCCT CCGGCTTTTT
```

- 601 ACCGAAGCGT GGCAGCATCT GTTGCGTAAA AATAAAAATC GGTTATCCTA
- 651 TGTGGATATG AGGTATAAGG ACGGATTTTC AGTCCGCTAT GCTTCCGACG 701 GTTTACCCGA AAAAGAATCC GAAGAATAG

This corresponds to the amino acid sequence <SEQ ID 2441; ORF 080>: m080.pep

- 1 MWDNAEAMER LTRWLLVMMA MLLAASGLVW FYNSNHLPVK QVSLKGNLVY
  - 51 SDKKTLGSLA KEYIHGNILR TDINGAQEAY RRYPWIASVM VRRRFPDTVE
  - 101 VVLTERKPVA RWGDHALVDG EGNVFEARLD RPGMPVFRGA EGTSAEMLRR 151 YDEFSTVLAK QGLGIKEMTY TARSAWIVVL DNGITVRLGR ENEMKRLRLF 201 TEAWQHLLRK NKNRLSYVDM RYKDGFSVRY ASDGLPEKES EE*

Computer analysis of this amino acid sequence gave the following results:

#### Homology with a predicted ORF from N. gonorrhoeae

ORF 080 shows 97.9% identity over a 242 aa overlap with a predicted ORF (ORF 080.ng) from *N. gonorrhoeae*:

m080/g080

m080.pep	10 MWDNAEAMERLTRWLI	1111111111			11111111:1	
080	MWDNAEAMERLTRWLI 10	ZOMMAMLLAAS	30	LPVKQVSLKG 40	NLVYSDKKAL 50	GSLA 60
m080.pep	70 KEYIHGNILRTDINGA            KEYIHGNILRTDINGA	QEAYRRYPWI	 ASVMVRRRFF	 DTVEVVLTER	 .KPVARWGDHA	1111
m020 non	70 130 EGNVFEARLDRPGMPV	80 140	90 150	160	110 170	180
m080.pep	EGNVFEARLDREGMEV 	1111111111	1111111111	1111111111		111
m080.pep	190 DNGITVRLGRENEMKR            DNGITVRLGRENEMKR 190		111111111	1111111111	11:1111	
m080.pep	EEX    EEYWEQVWDILRPGVG 250	NGSTQISISY 260	KGRRTMEQQX 270			

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 245>: pes.080a

1	ATGTGGGATA	ATGCCGAAGC	GATGGAACGG	CTGACGCGCT	GGCTGCTTGT
51	CATGATGGCG	ATGCTGCTTG	CTGCGTCCGG	GCTGGTTTGG	TTTTACAATT
101	CGAATCATCT	GCCCGTCAAG	CAGGTGTCGC	TGAAGGGCAA	CCTAGTTTAT
151	TCCGATAAGA	AAGCATTGGG	CAGTTTGGCG	AAAGAATACA	TCCATGGGAA
201	TATTTTGAGG	ACGGACATCA	ATGGCGCACA	GGAGGCCTAC	CGCCGGTATC
251	CGTGGATTGC	GTCGGTCATG	GTGCGCCGCC	GTTTTCCCGA	CACGGTTGAG
301	GTCGTCCTGA	CCGAGCGCAA	GCCGGTCGCG	CGTTGGGGCG	ACCATGCCTT
351	GGTGGACGGC	GAAGGCAATG	TTTTTGAAGC	CCGTTTGGAC	AGACCCGGAA
401	TGCCGGTATT	CAGAGGCGCG	GAAGGAACGT	CTGCCGAAAT	GCTCCGCCGT
451	TATGACGAAT	TTTCGACTGT	TTTGGCAAAA	CAGGGTTTGG	GCATCAAAGA
501	GATGACCTAT	ACGGCACGTT	CGGCGTGGAT	TGTCGTTTTG	GACAACGGCA

601

TCACCGTCAG GCTCGGACGG GAAAACGAGA TGAAACGCCT CCGGCTTTTT

```
ACCGAAGCGT GGCAACATCT GTTGCGTAAA AATAAAAATC GGTTATCCTA
          TGTGGATATG AGGTATAAGG ACGGATTTTC AGTCCGCTAT GCTCCCGACG
     651
          GTTTACCCGA AAAAGAATCC GAAGAATAG
 This corresponds to the amino acid sequence <SEQ ID 246; ORF 080.a>:
 a080.pep
          MWDNAEAMER LTRWLLVMMA MLLAASGLVW FYNSNHLPVK QVSLKGNLVY
      51
          SDKKALGSLA KEYIHGNILR TDINGAQEAY RRYPWIASVM VRRRFPDTVE
     101
          VVLTERKPVA RWGDHALVDG EGNVFEARLD RPGMPVFRGA EGTSAEMLRR
     151
          YDEFSTVLAK QGLGIKEMTY TARSAWIVVL DNGITVRLGR ENEMKRLRLF
          TEAWQHLLRK NKNRLSYVDM RYKDGFSVRY APDGLPEKES EE*
     201
m080/a080
             99.2% identity over a 242 aa overlap
                    10
                             20
                                       30
                                                40
                                                         50
            MWDNAEAMERLTRWLLVMMAMLLAASGLVWFYNSNHLPVKQVSLKGNLVYSDKKTLGSLA
m080.pep
             a080
            MWDNAEAMERLTRWLLVMMAMLLAASGLVWFYNSNHLPVKQVSLKGNLVYSDKKALGSLA
                             20
                                       30
                                                40
                                                         50
                                                                  60
                    70
                             80
                                       90
                                               100
                                                        110
m080.pep
            KEYIHGNILRTDINGAQEAYRRYPWIASVMVRRRFPDTVEVVLTERKPVARWGDHALVDG
            a080
            {	t KEYIHGNILRTDINGAQEAYRRYPWIASVMVRRRFPDTVEVVLTERKPVARWGDHALVDG}
                    70
                             80
                                      90
                                              100
                                                        110
                   130
                            140
                                     150
                                              160
                                                        170
                                                                 180
            {\tt EGNVFEARLDRPGMPVFRGAEGTSAEMLRRYDEFSTVLAKQGLGIKEMTYTARSAWIVVL}
m080.pep
            a080
            EGNVFEARLDRPGMPVFRGAEGTSAEMLRRYDEFSTVLAKQGLGIKEMTYTARSAWIVVL
                   130
                            140
                                     150
                                              160
                                                        170
                                                                 180
                  190
                            200
                                     210
                                              220
                                                        230
                                                                 240
            DNGITVRLGRENEMKRLRLFTEAWQHLLRKNKNRLSYVDMRYKDGFSVRYASDGLPEKES
m080.pep
            a080
            DNGITVRLGRENEMKRLRLFTEAWQHLLRKNKNRLSYVDMRYKDGFSVRYAPDGLPEKES
                  190
                            200
                                     210
                                              220
m080.pep
            EEX
            \Pi\Pi
a080
            EEX
The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 247>:
    g081.seq
              ATGAAACCAC TGGACCTAAA TTTCATCTGC CAAGCCCTCA AGCTTCCGAT
           1
              GCCGTCTGAA AACAAACCCG TGTCGCGCAT CGTAACCGAC AGCCGCGATA
          51
         101
              TTCGGGAAGG CGATGTGTTT TTCGCATTGG CGGGCGGGCG GTTTGACGCG
              CATGATTTTG TTGGAGGCGT ATTGTCTGCG GGCGCGGCGG CGGTTGTGGT
              TTCGCGCGAA GATTGCGCGG CTTTGGGCGG CGCGTTGAAA GTCGATGACA
         201
              CGCTTGCCGC GTTGCAAACG TTGGCGAAGG CGTGGCGCGA TAATGTGAAC
         251
              CCGTTTGTGT TCGGCATTAC CGGTTCGGGC GGCAAGACGA CGGTGAAGGA
         301
              GATGCTGGCT GCGGTATTGC GCCGCCGTTT CGGCGATGAT GCCGTTTCGG
         351
             CGACGGCAGG CAACTTCAAC AACCACAtcg gaTTGCCGCT GACTTTATTG
         401
             AAATtaaAcg aAAAACACCG CTATGCCGTG ATTGAAATGG GCATGAACCA
         451
             TTTTGGcgaa ctggcggtTt taacgcaaaT CGCCAAACCC GATGCCGCTT
         501
```

TGGtcaACAA CGCCCTGCGC GCCCATGTCG GATGCGGTTt cgacggagtg

GGCGATATTG CCAAAGcgaa aagcGAGATT TatgcagGct tATGTTCAGA

CGGCATGGCA CTGATTCCTC AAGAAGATGC CAATATGGCT GTCTTCAAAA CGGCAACGTT TAATTTGAAT ACGTGCACTT TCGGCGTCGA TAGCGGCGAT

GTCCGCGCGG AAAATATCGT GCTGAAACCT TTGTCGTGCG AATTTGATTT

551

601

701

751

. 15

```
801 GGTGTGCGGC GACGAGCGCA CTGCCGTGGT GCTGCCTGTT CCCGGCCGCC
               ACAATGTCCA CAACGCCGCC GCTGCCGCCG CGCTGGCTTT GGCTGCCGGT
               TTGAGTTTGA ACGATGTGGC GGAAGGTTTG CAAGGCTTCA GCAACATCAA
               AGGCCGTCTG AACGTCAAAG CCGGCATCAA GGGCGCAACC CTGATTGACG
         1001 ATACTTATAA TGCGAATCCC GACAGTATGA AAGCCGCGGT TGACGTGTTG
         1051 GCGCGTATGC CTGCGCCGCG CATTTTCGTG ATGGGCGATA TGGGCGAACT
         1101 GGGCGAGGAC gaAGCCGCCG CCATGCACGC CGAagtcgGC GCGTACGCCC
         1151 GCGACCAAGG CATCGAAGCG GCTTATTTTG TCGGCGACAA CAGCGTCGAA
         1201 GCGGcggaAA AATTTGGCGC GGACGGTTTG TGGTTCGCCG CCAAAGACCC
         1251 GTTGATTCAA GTGTTGAGCC ACGATTTGCC CGAACGCGCC ACCGTGTTGG
               TGAAAGGTTC GCGCTTTATG CAGAtggAAG AAGTGGTCGA GGCATTGGAG
               GATAAGTga
This corresponds to the amino acid sequence <SEQ ID 248; ORF 081.ng>:
     g081.pep
               MKPLDLNFIC QALKLPMPSE NKPVSRIVTD SRDIREGDVF FALAGGRFDA
           51 HDFVGGVLSA GAAAVVVSRE DCAALGGALK VDDTLAALQT LAKAWRDNVN
          101 PFVFGITGSG GKTTVKEMLA AVLRRRFGDD AVSATAGNFN NHIGLPLTLL
          151 KLNEKHRYAV IEMGMNHFGE LAVLTQIAKP DAALVNNALR AHVGCGFDGV
          201 GDIAKAKSEI YAGLCSDGMA LIPOEDANMA VFKTATFNLN TCTFGVDSGD
          251 VRAENIVLKP LSCEFDLVCG DERTAVVLPV PGRHNVHNAA AAAALALAAG
          301 LSLNDVAEGL QGFSNIKGRL NVKAGIKGAT LIDDTYNANP DSMKAAVDVL
          351 ARMPAPRIFV MGDMGELGED EAAAMHAEVG AYARDQGIEA AYFVGDNSVE
          401 AAEKFGADGL WFAAKDPLIQ VLSHDLPERA TVLVKGSRFM QMEEVVEALE
          451 DK*
The following partial DNA sequence was identified in N. meningitidis <SEO ID 249>:
     m081.seg
               ATGAAACCAC TGGACCTAAA TTTCATCTGC CAAGCCCTCA AGCTTCCGAT
           51 GCCGTCTGAA AGCAAACCCG TGTCGCGCAT CGTAACCGAC AGCCGCGACA
          101 TCCGCGCGGG CGATGTGTT TTCGCATTGG CGGGCGAGCG GTTTGACGCG
          151 CATGATTTG TTGAAGACGT ATTGGCTGCT GGTGCGGCGG CGGTTGTGGT
          201 TTCGCGCGAA GATTGTGCTG CAATGGATGG CGCGTTGAAA GTCGATGACA
              CGCTTGCCGC ATTGCAAACG CTGGCAAAGG CGTGGCGTGA AAATGTGAAT
               CCGTTTGTGT TCGGCATTAC CGGTTCGGGC GGCAAGACGA CGGTGAAGGA
               AATGCTGGCT GCGGTATTGC GCCGCCGTTT CGGCGATGAT GCCGTGTTGG
               CGACGGCAGG CAACTTCAAC AACCATATCG GATTGCCGCT GACTTTGTTG
          401
              AAGTTAAACG AAAAACACCG CTATGCCGTG ATTGAAATGG GCATGAACCA
          451
               TTTCGGCGAA CTGGCGGTTT TAACGCAMAT CGCCAAACCA AATGCCGCAT
          501
              TGGTCAACAA CGCCATGCGC GCCCATGTCG GCTGCGGTTT CGACGGAGTG
          551
               GGCGATATTG CCAAAGCGAA AAGCGAGATT TACCAAGGTT TATGTTCAGA
              CGGCATTGCA CTGATTCCTC AAGAAGATGC CAATATGGCT GTCTTCAAAA
          701
              CGGCAACGCT TAATTTGAAT ACGCGCACTT TCGGCATCGA TAGCGGCGAT
              GTTCACGCGG AAAATATTGT GCTGAAACCG TTGTCGTGCG AATTTGATTT
          801 GGTGTGCGGC GATGAGCGCG CCGCCGTGGT GCTGCCTGTT CCCGGCCGCC
          851 ACAATGTCCA CAACGCCGCC GCTGCCGCCG CGCTGGCTTT GGCTGCGGGT
               TTGAGTTTGA ACGATGTGGC GGAAGGTTTG AAAGGCTTCA GCAATATCAA
          901
          951 AGGCCGTCTG AACGTCAAAT CCGGAATCAA GGGCGCAACC CTGATTGACG
               ATACTTATAA TGCGAACCCT GACAGCATGA AAGCTGCGAT TGACGTGTTG
               GCGCGTATGC CTGCGCCGCG TATTTTCGTG ATGGGCGATA TGGGCGAACT
         1051
         1101
               GGGCGAACTG GGCGAGGACG AAGCCGCCGC TATGCACGCC GAAGTCGGCG
         1151
              CGTATGCCCG CGACCAAGGC ATCGAAGCGG CTTATTTTGT CGGCGACAAC
         1201 AGCGTCGAAG CGGCGGAAAA ATTTGGCGCG GACGGTTTGT GGTTCGCCGC
         1251 CAAAGACCCG TTGATTCAAG TGTTGCGCCA CGATTTGCCC GAACGCGCCA
         1301 CCGTGTTGGT GAAAGGTTCG CGCTTTATGC AGATGGAAGA AGTGGTCGAG
         1351 GCATTGGAGG ATAAGTGA
This corresponds to the amino acid sequence <SEQ ID 250; ORF 081>:
     m081.pep
            1
              MKPLDLNFIC QALKLPMPSE SKPVSRIVTD SRDIRAGDVF FALAGERFDA
           51 HDFVEDVLAA GAAAVVVSRE DCAAMDGALK VDDTLAALQT LAKAWRENVN
          101 PFVFGITGSG GKTTVKEMLA AVLRRRFGDD AVLATAGNFN NHIGLPLTLL
          151 KLNEKHRYAV IEMGMNHFGE LAVLTXIAKP NAALVNNAMR AHVGCGFDGV
```

201 GDIAKAKSEI YQGLCSDGIA LIPQEDANMA VFKTATLNLN TRTFGIDSGD
251 VHAENIVLKP LSCEFDLVCG DERAAVVLPV PGRHNVHNAA AAAALALAAG
301 LSLNDVAEGL KGFSNIKGRL NVKSGIKGAT LIDDTYNANP DSMKAAIDVL
351 ARMPAPRIFV MGDMGELGEL GEDEAAAMHA EVGAYARDQG IEAAYFVGDN
401 SVEAAEKFGA DGLWFAAKDP LIQVLRHDLP ERATVLVKGS RFMQMEEVVE
451 ALEDK*

Computer analysis of this amino acid sequence gave the following results: Homology with a predicted ORF from N. gonorrhoeae

ORF 081 shows 94.1% identity over a 455 aa overlap with a predicted ORF (ORF 081.ng) from N. gonorrhoeae:

m081/g081

m081.pep	10 MKPLDLNFICQAL           MKPLDLNFICQAL 10		 SRIVTDSRE	 IREGDVFFAL		1 11.1
m081.pep	70 GAAAVVVSREDCA	80 AMDGALKVDDT	30 90 LAALOTLAK	40 100 AWRENVNPFVF	50 110 GITGSGGKT	120 TVKEMIA
g081		:	1111111	111:111111		1111111
m081.pep	130 AVLRRRFGDDAVL	140 ATAGNFNNHIG	150 LPLTLLKLN	160 EKHRYAVIEMG 	170 MNHFGELAVI	180 LTXIAKP
g081	AVLRRRFGDDAVS2	ATAGNFNNHIG: 140	LPLTLLKLN 150	EKHRYAVIEMG 160	MNHFGELAVI 170	 LTQIAKP 180
m081.pep	190 NAALVNNAMRAHVO :      :			11111:1111	111111111	11.111
g081	DAALVNNALRAHVO 190 250	CGFDGVGDIAI 200	KAKSEIYAGI 210	LCSDGMALIPQ 220	EDANMAVFKI 230	PATFNLN 240
m081.pep	TRTFGIDSGDVHAE			:	111111111	11111
•	250	260 320	270 330	280 340	290 350	300
m081.pep	LSLNDVAEGLKGFS             LSLNDVAEGLQGFS	NIKGRLNVKSO	SIKGATLIDI	TYNANPDSMKI	AAIDVLARMP	TITLE.
m0.91 man	310 370	320 380	330 390	340 400	350 410	360
m081.pep	MGDMGELGELGEDE	1   1   1   1   1   1		 YFVGDNSVEA	 LEKFGADGLW	11111
m081.pep	430 LIQVLRHDLPERAT            LIQVLSHDLPERAT 420 430	440 VLVKGSRFMQM	450 EEVVEALED	11	410	

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 251>: a081.seq

509					
1	ATGAAACCAC	TGGACCTAAA	TTTCATCTGC	CAAGCCCTCA	AGCTTCCGAT
51	GCCGTCTGAA	AGCAAACCCG	TGTCGCGCAT	CGTAACCGAC	AGCCGCGACA
101	TCCGCGCGGG	CGATGTGTTT	TTCGCATTGG	CGGGCGGCG	GTTTGATGCG
151	CATGATTTTG	TTGAAGACGT	ATTGGCTGCG	GGTGCGGCGG	CGGTTGTGGT
201	TTCGCGCGAA	GATTGCGTTG	CAATGGATGG	CGCGTTGAAA	GTCGATGACA
251	CGCTTACCGC	GTTGCAAATG	TTGGCGAAGG	CGTGGCGCGA	GAATGTGAAC
301	CCGTTTGTGT	TCGGTATTAC	CGGCTCGGGC	GGCAAGACGA	CGGTGAAGGA
351	AATGTTGGCT	GCGGTATTGC	GCCGCCGTTT	CGGCGATAAT	GCCGTTTTGG
401	CGACGGCAGG	CAACTTCAAC	AACCACATCG	GATTGCCGTT	GACTTTGTTG
451	AAATTAAACG	AAAAACACCG	CTATGCCGTG	ATTGAAATGG	GTATGAACCA
501	TTTTGGCGAA	CTGGCGGTTT	TGACACAAAT	CGCCAAACCC	GATGCCGCAT
551	TGGTCAACAA	CGCCATGCGC	GCCCATGTCG	GCTGCGGTTT	CGACGGAGTG
601	GGCGATATTG	CCAAAGCGAA	AAGCGAGATT	TATCAAGGCT	TATGTTCAGA
651	CGGCATGGCG	CTGATTCCTC	AAGAAGATGC	CAATATGGCT	GTCTTCAAAA
701	CGGCAACGCT	TAATTTGAAT	ACGCGCACTT	TCGGCATCGA	TAGCGGCGAT
751	GTCCACGCGG	AAAATATCGT	GCTGAAACCG	TTGTCGTGCG	AATTTGATTT
801	GGTGTGCGGC	AACGAGTGCG	CAGCCGTGGT	TCTGCCCGTT	CCCGGCCGCC
851	ACAATGTCCA	CAACGCCGCC	GCCGCCGCCG	CGCTGTCTTT	GGCTGCAGGT
901	TTGAGTTTGA	ACGATGTGGC	GGAAGGTTTG	AAAGGCTTCA	GCAATATCAA
951	AGGCCGTCTG	AACGTCAAAT	CCGGAATCAA	GGGCGCAACC	CTGATTGACG
1001	ATACTTATAA	TGCGAACCCT	GACAGCATGA	AAGCTGCGGT	TGACGTGTTG
1051	GCGCGTATGC	CTGCGCCGCG	TATTTTCGTG	ATGGGCGATA	TGGGCGAACT
1101	GGGTGAGGAC	GAAGCCGCCG	CCATGCACGC	CGAAGTCGGC	GCGTACGCCC
1151	GCGACCAAGG	CATCGAAGCG	GCTTATTTTG	TCGGCGACAA	CAGCGTCGAA
1201	GCGGCGGAAA	AATTTGGCGC	GGACGGTTTG	TGGTTCGCCG	CCAAAGACCC
1251	GTTGATTCAA	GTGTTGCGCC	ACGATTTGCC	CGAACGCGCC	ACCGTGTTGG
1301	TGAAAGGTTC	GCGCTTTATG	CAGATGGAAG	AAGTGGTCGA	GGCATTGGAG
1351	GATAAGTGA				

### This corresponds to the amino acid sequence <SEQ ID 252; ORF 081.a>: a081.pep

F - F					
1	MKPLDLNFIC	QALKLPMPSE	SKPVSRIVTD	SRDIRAGDVF	FALAGGRFDA
51	HDFVEDVLAA	GAAAVVVSRE	DCVAMDGALK	VDDTLTALQM	LAKAWRENVN
101	PFVFGITGSG	GKTTVKEMLA	AVLRRRFGDN	AVLATAGNFN	NHIGLPLTLL
151	KLNEKHRYAV	IEMGMNHFGE	LAVLTQIAKP	DAALVNNAMR	AHVGCGFDGV
201	GDIAKAKSEI	YQGLCSDGMA	LIPQEDANMA	VFKTATLNLN	TRTFGIDSGD
251	VHAENIVLKP	LSCEFDLVCG	NECAAVVLPV	PGRHNVHNAA	AAAALSLAAG
301	LSLNDVAEGL	KGFSNIKGRL	NVKSGIKGAT	LIDDTYNANP	DSMKAAVDVL
351	ARMPAPRIFV	MGDMGELGED	EAAAMHAEVG	AYARDQGIEA	AYFVGDNSVE
401	AAEKFGADGL	WFAAKDPLIQ	VLRHDLPERA	TVLVKGSRFM	QMEEVVEALE
451	DK*				

#### m081/a081 96.7% identity over a 455 aa overlap

	10	20	30	40	50	60
m081.pep	MKPLDLNFICQALK	LPMPSESKPV	SRIVTDSRDI	RAGDVFFALA	GERFDAHDFV	EDVLAA
			111111111	HILLIIII	1 11111111	111111
a081	MKPLDLNFICQALK	LPMPSESKPV	'SRIVTDSRDI	RAGDVFFALA	GGRFDAHDFV	EDVLAA
	10	20	30	40	50	60
	70	80	90	100	110	120
m081.pep	GAAAVVVSREDCAAI	1DGALKVDDT	'LAALQTLAKA'	WRENVNPFVF	GITGSGGKTT	VKEMLA
	11111111111111		1:111 1111	1111111111	111111111	11111
a081	GAAAVVVSREDCVA	1DGALKVDDT	LTALQMLAKA	WRENVNPFVF	GITGSGGKTT	VKEMLA
	70	80	90	100	110	120
	130	140	150	160	170	180
m081.pep	AVLRRRFGDDAVLA	<b>FAGNFNNHIG</b>	LPLTLLKLNE	KHRYAVIEMG	MNHFGELAVL	TXIAKP
			111111111	HILLIHILL	11111111111	1 1111
a081	AVLRRRFGDNAVLA	ragnfnnhig	LPLTLLKLNE	KHRYAVIEMG	MNHFGELAVL	TQIAKP

268

	130	140	150	160	170	180
	190	200	010			
m081.pep		200 AHVGCGFDGVGD:	210 TAKAKSETYOGI	220	230	240
	: [ ]   [ ]   [ ]   [	11111111111		1111:1111		HILL
a081	DAALVNNAMR	AHVGCGFDGVGD:	IAKAKSEIYQGI	CSDGMALIPQ	EDANMAVFKTA	TLNLN
	190	200	210	220	230	240
	250	260	270	280	290	300
m081.pep	TRTFGIDSGD	JHAENIVLKPLS	CEFDLVCGDERA	AVVLPVPGRH	.TAAAAAHVN	AT.AAG
- 001				1111111111	11111111111	. 1 1 1 1
a081	TRTFGIDSGDV 250	/HAENIVLKPLS( 260	CEFDLVCGNECA 270			
	250	260	270	280	290	300
	310	320	. 330	340	350	360
m081.pep	LSLNDVAEGLE	KGFSNIKGRLNVF	KSGIKGATLIDD	TYNANPDSMK	AAIDVLARMPA	PRIFV
a081	ISINDVARCIA	CECNIKODINA		11111111		1111
4001	310	GFSNIKGRLNV 320	330	TYNANPDSMKA 340	AAVDVLARMPAI 350	
	510	320	550	340	330	360
	370	380	390	400	410	420
m081.pep	MGDMGELGELG	EDEAAAMHAEVG	AYARDQGIEAA	YFVGDNSVEA	EKFGADGLWF#	AAKDP
a081	MGDMGELGE	TITITITE DEAAAMHAEVG		VETICONSTIEN	EKECARCIUS	1111
	3	70 380	390	1F VGDNS VEAL 400	AERIGADGLWF7 410	AKDP
				200	440	
m081.pep	430	440	450			
mod1.pep		RATVLVKGSRFM				
a081		RATVLVKGSRFM				
		30 440				
The following	partial DNA	sequence wa	s identified i	n <i>N. gonorr</i>	hoeae <seo< td=""><td>ID 253&gt;:</td></seo<>	ID 253>:
g082.sec	I					
]	. aTGTGGTTG	T TGAAGTTGCC	TGCCGTCGCC	GAAACGGCA:	CATCGCCGAZ	<b>Y</b>
53 103	. ACGGCGGCG	C AATACCGCAG	CCAGCATCTC	CTTCACCGT	GTCTTGCCG	3
151	TTCCCCAAC	T AATGCCGAAC G TTTGCAACGC	GGCAAGCGGGT	TCACATTATO	GCGCCACGC	3
201		o illounicoc				
	GCCCAAAGC	C GCGCAATCTT	CGCGCGAAAC	CACAACCGCC	COCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	
251	. GCCCAAAGC	C GCGCAATCTT	CGCGCGAAAC	CACAACCGCC	GCCGCGCCC	}
251 301	GCCCAAAGC CAGACAATA AATGCGAAA	C GCGCAATCTT C GCCTCCAACA A ACACATCGCC	CGCGCGAAAC AAATCATGCG TTCCCGAATA	CACAACCGCC CGTCAAACCC TCGCGGCTGT	GCCGCGCCCGCCGCCCCCCCCCCCCCCCCCCCCCCCCC	<del>}</del>
301 351	GCCCAAAGC CAGACAATA AATGCGAAA GCGCGACAC	C GCGCAATCTT C GCCTCCAACA A ACACATCGCC G GGTTTGTTTT	CGCGCGAAAC AAATCATGCG TTCCCGAATA CAGACGGCAT	CACAACCGCC CGTCAAACCC TCGCGGCTGT CGGAAGCTTC	GCCGCGCCCGCGCCGCCCCCCCCCCCCCCCCCCCCCCC	
301 351 401	CAGACAAAGC CAGACAATA AATGCGAAA GCGCGACAC AGATGAAAT	C GCGCAATCTT C GCCTCCAACA A ACACATCGCC G GGTTTGTTT T TAGGTCCAGT	CGCGCGAAAC AAATCATGCG TTCCCGAATA CAGACGGCAT GGTTTCATAT	CACAACCGCC CGTCAAACCC TCGCGGCTGT CGGAAGCTTC TTGCTTTCGT	GCCGCGCCCGGCGGCGGCGCGCGCGGGGGCTTGGGGCTTGGGGCTTGGGGCTTGGGGCTTGGGGCTTGGGGGCTTGGGGGCTTGGGGGG	
301 351 401 451	GCCCAAAGC CAGACAATA AATGCGAAA GCGCGACAC AGATGAAAT GCGGCGGACAC	C GCGCAATCTT C GCCTCCAACA A ACACATCGCC G GGTTTGTTT T TAGGTCCAGT A CATCGGTAGC	CGCGCGAAAC AAATCATGCG TTCCCGAATA CAGACGGCAT GGTTTCATAT GGCTGATTTT	CACAACCGCC CGTCAAACCC TCGCGGCTGT CGGAAGCTTC TTGCTTTCGT TTTATCGCCT	GCCGCGCCCCGCGCGCCGCCGCCGCCGCGCGCGCGCGC	<del>)</del> 2 2 3
301 351 401 451 501	GCCCAAAGO CAGACAATA AATGCGAAA GCGCGACAC AGATGAAAT GCGGCGGAC GGTAAAACA	C GCGCAATCTT C GCCTCCAACA A ACACATCGCC G GGTTTGTTT T TAGGTCCAGT A CATCGGTAGC C AGATTATTTT	CGCGCGAAAC AAATCATGCG TTCCCGAATA CAGACGGCAT GGTTTCATAT GGCTGATTTT CCCATTCTCA	CACAACCGCC CGTCAAACCC TCGCGGCTGT CGGAAGCTTC TTGCTTTCGT TTTATCGCCT TTCGGCATTT	GCCGCGCCCCGCGCCGCCGCCGCCCGCCGCCGCCGCCGC	
301 351 401 451 501 551	GCCCAAAGC CAGACAATA AATGCGAAA GCGCGACAC AGATGAAAT GCGGCGGAC GGTAAAACA TATCATTTT	C GCGCAATCTT C GCCTCCAACA A ACACATCGCC G GGTTTGTTT T TAGGTCCAGT A CATCGGTAGC C AGATTATTTT T TAGACGTATT	CGCGCGAAAC AAATCATGCG TTCCCGAATA CAGACGGCAT GGTTTCATAT GGCTGATTTT CCCATTCTCA TTTAGCCGAT	CACAACCGCC CGTCAAACCC TCGCGGCTGT CGGAAGCTTC TTGCTTTCGT TTTATCGCCT TTCGGCATTT TTGCCTTTTC	GCCGCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	
301 351 401 451 501 551	GCCCAAAGO CAGACAATA AATGCGAAA GCGCGACAC AGATGAAAT GCGGCGGAC GGTAAAACA TATCATTTT CGGCGCGGG	C GCGCAATCTT C GCCTCCAACA A ACACATCGCC G GGTTTGTTT T TAGGTCCAGT A CATCGGTAGC C AGATTATTT T TAGACGTATT G TCGTCGGACT	CGCGCGAAAC AAATCATGCG TTCCCGAATA CAGACGGCAT GGTTTCATAT GGCTGATTTT CCCATTCTCA TTTAGCCGAT GTCTGTCGAT	CACAACCGCC CGTCAAACCC TCGCGGCTGT CGGAAGCTTC TTGCTTTCGT TTTATCGCCT TTCGGCATTT TTGCCTTTTC AAAGGCAAGG	GCCGCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	
301 351 401 451 501 551 601 651 701	GCCCAAAGO CAGACAATA AATGCGAAA GCGCGACAC AGATGAAAT GCGGCGGAC GGTAAAACA TATCATTTT CGGCGCGGC CGCCCGGCA	C GCGCAATCTT C GCCTCCAACA A ACACATCGCC G GGTTTGTTTT T TAGGTCCAGT A CATCGGTAGC C AGATTATTTT T TAGACGTATT T TAGACGTATT G TCGTCGGACA C ATCGGGGACA G TTTCGATACC	CGCGCGAAAC AAATCATGCG TTCCCGAATA CAGACGGCAT GGTTTCATAT GGCTGATTTT CCCATTCTCA TTTAGCCGAT GTCTGTCGAT TTCCCCCAAA CGTCCAACTG	CACAACCGCC CGTCAAACCC TCGCGGCTGT TCGCAAGCTTC TTGCTTTCGT TTTATCGCCT TTCGGCATTT TTGCCTTTTC AAAGGCAAGC AATCATAGCC CCGAATCCGC	GCCGCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	
301 351 401 451 501 551 601 651 701 This correspon	GCCCAAAGO CAGACAATA AATGCGAAA GCGCGACAC AGATGAAAT GCGGCGGAC GGTAAAACA TATCATTTT CGGCGCGGC CGCCCGGCA AACTCGTCG ds to the am:	C GCGCAATCTT C GCCTCCAACA A ACACATCGCC G GGTTTGTTTT T TAGGTCCAGT A CATCGGTAGC C AGATTATTTT T TAGACGTATT T TAGACGTATT G TCGTCGGACA C ATCGGGGACA G TTTCGATACC	CGCGCGAAAC AAATCATGCG TTCCCGAATA CAGACGGCAT GGTTTCATAT GGCTGATTTT CCCATTCTCA TTTAGCCGAT GTCTGTCGAT TTCCCCCAAA CGTCCAACTG	CACAACCGCC CGTCAAACCC TCGCGGCTGT TCGCAAGCTTC TTGCTTTCGT TTTATCGCCT TTCGGCATTT TTGCCTTTTC AAAGGCAAGC AATCATAGCC CCGAATCCGC	GCCGCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	
301 351 401 451 501 551 601 653 701 This correspon g082.pep	GCCCAAAGO CAGACAATA AATGCGAAA GCGCGACAC AGATGAAAT GCGGCGGAC GGTAAAACA TATCATTTT CGGCGCGGC CGCCCGGCA AACTCGTCG ds to the am:	C GCGCAATCTT C GCCTCCAACA A ACACATCGCC G GGTTTGTTTT T TAGGTCCAGT A CATCGGTAGC C AGATTATTTT T TAGACGTATT G TCGTCGGACT C ATCGGGGACA G TTTCGATACC ino acid seque	CGCGCGAAAC AAATCATGCG TTCCCGAATA CAGACGGCAT GGTTTCATAT GGCTGATTTT CCCATTCTCA TTTAGCCGAT GTCTGTCGAT TTCCCCCAAA CGTCCAACTG CCCCCAACTG	CACAACCGCC CGTCAAACCC TCGCGGCTGT CGGAAGCTTC TTGCTTTCGT TTTATCGCCT TTGCCTTTTC AAAGGCAAGC AATCATAGCC CCGAATCCGC D 254; ORF	GCCGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	
301 351 401 451 501 551 601 651 701 This correspon g082.pep	GCCCAAAGC CAGACAATA AATGCGAAA GCGCGACAC AGATGAAAT GCGGCGGAC GGTAAAACA TATCATTTT CGGCGCGGC AACTCGTCG ds to the am:	C GCGCAATCTT C GCCTCCAACA A ACACATCGCC G GGTTTGTTTT T TAGGTCCAGT A CATCGGTAGC C AGATTATTTT T TAGACGTATT G TCGTCGGACT C ATCGGGGACA G TTTCGATACC ino acid seque	CGCGCGAAAC AAATCATGCG TTCCCGAATA CAGACGGCAT GGTTTCATAT GGCTGATTTT CCCATTCTCA TTTAGCCGAT GTCTGTCGAT TTCCCCCAAA CGTCCAACTG CCCCCAACTG CCCCCCAACTG CCCCCAACTG CCCCCCAACTG CCCCCCAACTG CCCCCCAACTG CCCCCCAACTG CCCCCCAACTG CCCCCCAACTG CCCCCCAACTG CCCCCCCAACTG CCCCCCAACTG CCCCCCAACTG CCCCCCAACTG CCCCCCAACTG CCCCCCAACTG CCCCCCCAACTG CCCCCCCCCC	CACAACCGCC CGTCAAACCC TCGCGGCTGT CGGAAGCTTC TTGCTTTCGT TTTATCGCCT TTGCCTTTTC AAAGGCAAGC AATCATAGCC CCGAATCCGC D 254; ORF	GCCGCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	
301 351 401 451 501 551 601 651 701 This correspon g082.pep	GCCCAAAGC CAGACAATA AATGCGAAA GCGCGACAC AGATGAAAT GCGGCGGAC GGTAAAACA TATCATTTT CGGCGCGGC AACTCGTCG ds to the am: MWLLKLPAV FANVCNAAS	C GCGCAATCTT C GCCTCCAACA A ACACATCGCC G GGTTTGTTTT T TAGGTCCAGT A CATCGGTAGC C AGATTATTTT T TAGACGTATT G TCGTCGGACT C ATCGGGGACA G TTTCGATACC ino acid seque A ETASSPKRRR V SSTFNAPPKA	CGCGCGAAAC AAATCATGCG TTCCCGAATA CAGACGGCAT GGTTTCATAT GGCTGATTTT CCCATTCTCA TTTAGCCGAT GTCTGTCGAT TTCCCCCAAA CGTCCAACTG CCCCAACTG CCCCCAACTG CCCCCCAACTG CCCCCAACTG CCCCCAACTG CCCCCAACTG CCCCCAACTG CCCCCAACTG CCCCCAACTG CCCCCAACTG CCCCCAACTG CCCCCAACTG CCCCCCAACTG CCCCCAACTG CCCCCCAACTG CCCCCCCCCC	CACAACCGCC CGTCAAACCC TCGCGGCTGT TCGCGAAGCTTC TTTATCGCCT TTCGGCATTT TTGCCTTTTC AAAGGCAAGC AATCATAGCC CCGAATCCGC D 254; ORF VLPPEPVMPN AAPADNTPPT	GCCGCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	
301 351 401 451 501 551 601 653 701 This correspon g082.pep	GCCCAAAGC CAGACAATA AATGCGAAA GCGCGACAC AGATGAAAT GCGGCGGAC GGTAAAACA TATCATTTT CGGCGCGGC AACTCGTCG ds to the am: MWLLKLPAV FANVCNAAS NAKNTSPSR	C GCGCAATCTT C GCCTCCAACA A ACACATCGCC G GGTTTGTTTT T TAGGTCCAGT A CATCGGTAGC C AGATTATTTT T TAGACGTATT G TCGTCGGACT C ATCGGGGACA G TTTCGATACC ino acid seque A ETASSPKRRR V SSTFNAPPKA I SRLSVTMRDT	CGCGCGAAAC AAATCATGCG TTCCCGAATA CAGACGGCAT GGTTTCATAT GGCTGATTTT CCCATTCTCA TTTAGCCGAT TTCCCCCAAA CGTCCAACTG CCCCAACTG CCCCAACTG CCCCAACTG CCCCAACTG CCCCAACTG CCCCAACTG CCCCAACTG CCCCAACTG CCCCCAACTG CCCCCCAACTG CCCCCCAACTG CCCCCAACTG CCCCCCAACTG CCCCCCCAACTG CCCCCCAACTG CCCCCCAACTG CCCCCCAACTG CCCCCCAACTG CCCCCCAACTG CCCCCCAACTG CCCCCCAACTG CCCCCCAACTG CCCCCCAACTG CCCCCAACTG CCCCCAACTG CCCCCAACTG CCCCCAACTG CCCCCCAACTG CCCCCAACTG CCCCCAACTG CCCCCCAACTG CCCCCCCAACTG CCCCCCCAACTG CCCCCCCCCC	CACAACCGCC CGTCAAACCC TCGCGGCTGT TCGCGAAGCTTC TTTATCGCCT TTTGCCTTTTC AAAGGCAAGC AATCATAGCC CCGAATCCGC D 254; ORF VLPPEPVMPN AAPADNTPPT RAWOMKFRSS	GCCGCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	
301 351 401 451 501 551 601 653 701 This correspon g082.pep	GCCCAAAGC CAGACAATA AATGCGAAA GCGCGACAC AGATGAAAT GCGGCGGGAC GGTAAAACA TATCATTTT CGGCGCGGCA AACTCGTCG ds to the am: MWLLKLPAV FANVCNAAS NAKNTSPSR AADTSVAAD	C GCGCAATCTT C GCCTCCAACA A ACACATCGCC G GGTTTGTTTT T TAGGTCCAGT A CATCGGTAGC C AGATTATTTT T TAGACGTATT G TCGTCGGACT C ATCGGGGACA G TTTCGATACC ino acid seque A ETASSPKRRR V SSTFNAPPKA I SRLSVTMRDT F FIACFAVVKH	CGCGCGAAAC AAATCATGCG TTCCCGAATA CAGACGGCAT GGTTTCATAT GGCTGATTTT CCCATTCTCA TTTAGCCGAT TTCCCCCAAA CGTCCAACTG CCCCAACTG CCCCAACTG CCCCAACTG CCCCAACTG CCCCCAACTG  CCCCCAAACTG CCCCCAAACTG CCCCCAAACTG CCCCCAAACTG CCCCCAAACTG CCCCCAACTG CCCCCCAACTG CCCCCCCAACTG CCCCCCAACTG CCCCCCCAACTG CCCCCCCAACTG CCCCCCCAACTG CCCCCCCCCC	CACAACCGCC CGTCAAACCC TCGCGGCTGT TCGCGAAGCTTC TTGCTTTCGT TTTATCGCCT TTGCCTTTTC AAAGGCAAGC AATCATAGCC CCGAATCCGC D 254; ORF VLPPEPVMPN AAPADNTPPT RAWQMKFRSS FLYVSFFRRI	GCCGCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	
301 351 401 451 501 551 601 653 701 This correspon g082.pep 1 51 101 151 201	GCCCAAAGC CAGACAATA AATGCGAAA GCGCGACAC AGATGAAAT GCGGCGGGAC GGTAAAACA TATCATTTT CGGCGCGGCA AACTCGTCG ds to the am: MWLLKLPAV FANVCNAAS NAKNTSPSR AADTSVAAD RRGVVGLSV	C GCGCAATCTT C GCCTCCAACA A ACACATCGCC G GGTTTGTTTT T TAGGTCCAGT A CATCGGTAGC C AGATTATTTT T TAGACGTATT G TCGTCGGACT C ATCGGGGACA G TTTCGATACC ino acid seque A ETASSPKRRR V SSTFNAPPKA I SRLSVTMRDT F FIACFAVVKH D KGKVIAFARH	CGCGCGAAAC AAATCATGCG TTCCCGAATA CAGACGGCAT GGTTTCATAT GGCTGATTTT CCCATTCTCA TTTAGCCGAT TTCCCCCAAA CGTCCAACTG Ence <seq aqssrettta="" glfsdgigsl="" igdippkiia<="" ii="" ntaasisftv="" rlfshshsaf="" td=""><td>CACAACCGCC CGTCAAACCC CGTCAAACCC TCGCGGCTGT CGGAAGCTTC TTGCTTTCGT TTTATCGCCT TTGCCTTTTC AAAGGCAAGC AATCATAGCC CCGAATCCGC D 254; ORF VLPPEPVMPN AAPADNTPPT RAWQMKFRSS FLYVSFFRRI VIGOLVGFDT</td><td>GCCGCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC</td><td></td></seq>	CACAACCGCC CGTCAAACCC CGTCAAACCC TCGCGGCTGT CGGAAGCTTC TTGCTTTCGT TTTATCGCCT TTGCCTTTTC AAAGGCAAGC AATCATAGCC CCGAATCCGC D 254; ORF VLPPEPVMPN AAPADNTPPT RAWQMKFRSS FLYVSFFRRI VIGOLVGFDT	GCCGCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	
301 351 401 451 501 551 601 653 701 This correspon g082.pep 1 51	GCCCAAAGO CAGACAATA AATGCGAAA GCGCGACAC AGATGAAAT GCGGCGGGAC GGTAAAACA TATCATTTT CGGCGCGGCA AACTCGTCG ds to the am: MWLLKLPAV FANVCNAAS NAKNTSPSR AADTSVAAD RRGVVGLSV partial DNA	C GCGCAATCTT C GCCTCCAACA A ACACATCGCC G GGTTTGTTTT T TAGGTCCAGT A CATCGGTAGC C AGATTATTTT T TAGACGTATT G TCGTCGGACT C ATCGGGGACA G TTTCGATACC ino acid seque A ETASSPKRRR V SSTFNAPPKA I SRLSVTMRDT F FIACFAVVKH D KGKVIAFARH	CGCGCGAAAC AAATCATGCG TTCCCGAATA CAGACGGCAT GGTTTCATAT GGCTGATTTT CCCATTCTCA TTTAGCCGAT TTCCCCCAAA CGTCCAACTG Ence <seq aqssrettta="" glfsdgigsl="" igdippkiia<="" ii="" ntaasisftv="" rlfshshsaf="" td=""><td>CACAACCGCC CGTCAAACCC CGTCAAACCC TCGCGGCTGT CGGAAGCTTC TTGCTTTCGT TTTATCGCCT TTGCCTTTTC AAAGGCAAGC AATCATAGCC CCGAATCCGC D 254; ORF VLPPEPVMPN AAPADNTPPT RAWQMKFRSS FLYVSFFRRI VIGOLVGFDT</td><td>GCCGCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC</td><td></td></seq>	CACAACCGCC CGTCAAACCC CGTCAAACCC TCGCGGCTGT CGGAAGCTTC TTGCTTTCGT TTTATCGCCT TTGCCTTTTC AAAGGCAAGC AATCATAGCC CCGAATCCGC D 254; ORF VLPPEPVMPN AAPADNTPPT RAWQMKFRSS FLYVSFFRRI VIGOLVGFDT	GCCGCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	
301 351 401 451 501 551 601 651 701 This correspon g082.pep 1 51 101 151 201 The following m082.seq	GCCCAAAGO CAGACAATA AATGCGAAA GCGCGACAC AGATGAAAT GCGGCGGAC GGTAAAACA TATCATTTT CGGCGCGGC AACTCGTCG ds to the am: MWLLKLPAV FANVCNAAS NAKNTSPSR AADTSVAAD RRGVVGLSV. partial DNA	C GCGCAATCTT C GCCTCCAACA A ACACATCGCC G GGTTTGTTTT T TAGGTCCAGT A CATCGGTAGC C AGATTATTTT T TAGACGTATT G TCGTCGGACT C ATCGGGGACA G TTTCGATACC ino acid seque A ETASSPKRRR V SSTFNAPPKA I SRLSVTMRDT F FIACFAVVKH D KGKVIAFARH Sequence Was T TGAAGTTGCC	CGCGCGAAAC AAATCATGCG TTCCCGAATA CAGACGGCAT GGTTTCATAT GGCTGATTTT CCCATTCTCA TTTAGCCGAT TTCCCCCAAA CGTCCAACTG ENCE <seq aqssrettta="" glfsdgigsl="" identified="" igdippkiia="" ii="" in<="" ntaasisftv="" rlfshshsaf="" s="" td=""><td>CACAACCGCC CGTCAAACCG CGTCAAACCG TCGCGGCTGT CGGAAGCTTC TTGCTTTCGT TTTATCGCCT TTGCCTTTTC AAAGGCAAGC AATCATAGCC CCGAATCCGC D 254; ORF VLPPEPVMPN AAPADNTPPT RAWQMKFRSS FLYVSFFRRI VIGQLVGFDT N. mening. AACACCGCAT</td><td>GCCGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG</td><td>ID 255&gt;:</td></seq>	CACAACCGCC CGTCAAACCG CGTCAAACCG TCGCGGCTGT CGGAAGCTTC TTGCTTTCGT TTTATCGCCT TTGCCTTTTC AAAGGCAAGC AATCATAGCC CCGAATCCGC D 254; ORF VLPPEPVMPN AAPADNTPPT RAWQMKFRSS FLYVSFFRRI VIGQLVGFDT N. mening. AACACCGCAT	GCCGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	ID 255>:
301 351 401 451 501 551 601 651 701 This correspon g082.pep 1 51 201 The following m082.seq 1 51	GCCCAAAGO CAGACAATA AATGCGAAA GCGCGACAC AGATGAAAT GCGGCGGACA GGTAAAACA TATCATTTT CGGCGCGGCA AACTCGTCG ds to the am: MWLLKLPAV FANVCNAAS NAKNTSPSR AADTSVAAD RRGVVGLSV partial DNA ATGNNGTTG' ACGGCGGCG	C GCGCAATCTT C GCCTCCAACA A ACACATCGCC G GGTTTGTTTT T TAGGTCCAGT A CATCGGTAGC C AGATTATTTT T TAGACGTATT G TCGTCGGACT C ATCGGGGACA G TTTCGATACC ino acid seque A ETASSPKRRR V SSTFNAPPKA I SRLSVTMRDT F FIACFAVVKH D KGKVIAFARH SEQUENCE WAS I TGAAGTTGCC C AATACCGCAG	CGCGCGAAAC AAATCATGCG TTCCCGAATA CAGACGGCAT GGTTTCATAT GGCTGATTTT CCCATTCTCA TTTAGCCGAT TTCCCCCAAA CGTCCAACTG ENCE <seq aqssrettta="" ccagcattcc<="" glfsdgigsl="" identified="" igdippkiia="" ii="" in="" ntaasisftv="" rlfshshsaf="" s="" td="" tgccgtcgcc=""><td>CACAACCGCC CGTCAAACCGC CGTCAAACCC TCGCGGCTGT CGGAAGCTTC TTGCTTTCGT TTTATCGCCT TTGCCTTTTC AAAGGCAAGC AATCATAGCC CCGAATCCGC D 254; ORF VLPPEPVMPN AAPADNTPPT RAWQMKFRSS FLYVSFFRRI VIGQLVGFDT N. mening AACACCGCAT</td><td>GCCGCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC</td><td>ID 255&gt;:</td></seq>	CACAACCGCC CGTCAAACCGC CGTCAAACCC TCGCGGCTGT CGGAAGCTTC TTGCTTTCGT TTTATCGCCT TTGCCTTTTC AAAGGCAAGC AATCATAGCC CCGAATCCGC D 254; ORF VLPPEPVMPN AAPADNTPPT RAWQMKFRSS FLYVSFFRRI VIGQLVGFDT N. mening AACACCGCAT	GCCGCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	ID 255>:
301 351 401 451 501 551 601 651 701 This correspon g082.pep 1 51 101 The following m082.seq 1 51 101	GCCCAAAGO CAGACAATA AATGCGAAA GCGCGACAC AGATGAAAT GCGGCGGACA GGTAAAACA TATCATTTT CGGCGCGGCA AACTCGTCG ds to the am: MWLLKLPAV FANVCNAAS NAKNTSPSR AADTSVAAD RRGVVGLSV partial DNA ATGNNGTTG' ACGGCGGCGC	C GCGCAATCTT C GCCTCCAACA A ACACATCGCC G GGTTTGTTTT T TAGGTCCAGT A CATCGGTAGC C AGATTATTTT T TAGACGTATT G TCGTCGGACT C ATCGGGGACA G TTTCGATACC ino acid seque A ETASSPKRRR V SSTFNAPPKA I SRLSVTMRDT F FIACFAVVKH D KGKVIAFARH SEQUENCE WAS I TGAAGTTGCC C AATACCGCAG I TGAAGTTGCC C AATACCGCAG I AATGCCGAAC	CGCGCGAAAC AAATCATGCG TTCCCGAATA CAGACGGCAT GGTTTCATAT GGCTGATTTT CCCATTCTCA TTTAGCCGAT TTCCCCCAAA CGTCCAACTG CCCCAACTG CCCCAACTG CCCCAACTG CCCCAACTG CCCCCAACTG CCCCCAACTG CCCCCCCCCC	CACAACCGCC CGTCAAACCGC CGTCAAACCC TCGCGGCTGT CGGAAGCTTC TTGCTTTCGT TTTATCGCCT TTGCCTTTTC AAAGGCAAGC AATCATAGCC CCGAATCCGC D 254; ORF VLPPEPVMPN AAPADNTPPT RAWQMKFRSS FLYVSFFRRI VIGQLVGFDT N. mening AACACCGCAT CTTCACCGTC TCACATTTTC	GCCGCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	ID 255>:
301 351 401 451 501 551 601 651 701 This correspon g082.pep 1 51 201 The following m082.seq 1 51	GCCCAAAGO CAGACAATA AATGCGAAA GCGCGACAC AGATGAAAT GCGGCGGACA GGTAAAACA TATCATTTT CGGCGCGGCA AACTCGTCG ds to the am: MWLLKLPAV FANVCNAAS NAKNTSPSR AADTSVAAD RRGVVGLSV partial DNA ATGNNGTTG' ACGGCGGCGC	C GCGCAATCTT C GCCTCCAACA A ACACATCGCC G GGTTTGTTTT T TAGGTCCAGT A CATCGGTAGC C AGATTATTTT T TAGACGTATT G TCGTCGGACT C ATCGGGGACA G TTTCGATACC ino acid seque A ETASSPKRRR V SSTFNAPPKA I SRLSVTMRDT F FIACFAVVKH D KGKVIAFARH SEQUENCE WAS I TGAAGTTGCC C AATACCGCAG	CGCGCGAAAC AAATCATGCG TTCCCGAATA CAGACGGCAT GGTTTCATAT GGCTGATTTT CCCATTCTCA TTTAGCCGAT TTCCCCCAAA CGTCCAACTG CCCCAACTG CCCCAACTG CCCCAACTG CCCCAACTG CCCCCAACTG CCCCCAACTG CCCCCCCCCC	CACAACCGCC CGTCAAACCGC CGTCAAACCC TCGCGGCTGT CGGAAGCTTC TTGCTTTCGT TTTATCGCCT TTGCCTTTTC AAAGGCAAGC AATCATAGCC CCGAATCCGC D 254; ORF VLPPEPVMPN AAPADNTPPT RAWQMKFRSS FLYVSFFRRI VIGQLVGFDT N. mening AACACCGCAT CTTCACCGTC TCACATTTTC	GCCGCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	ID 255>:

201 ATCCATTGCA GCACAATCTT CGCGCGAAAC CACAACCGCC GCCGCACCAG 251 CAGCCAATAC GTCTTCAACA AAATCATGCG CGTCAAACCG CTCGCCCGCC 301 AATGCGAAAA ACACATCGCC CGCGCGGATG TCGCGGCTGT CGGTTACGAT 351 GCGCGACACG GGTTTGCTTT CAGACGGCAT CGGAAGCTTG AGGGCTTGGC 401 AGATGAAATT TAGGTCCAGT GGTTTCATAT TTACTTTCGT TAATATTCGG 451 GCGGCGGACA CATCGGTAGC GGCTGATTTT TTTATCGCCT GTTTTGCTGT 501 GGTAAAACAC AGATTATTTT CCCATTCTCA TTCGGsATTT TTTCTGTACG 551 TATCATTTT TAGACGTATT TTTAGTCGAT TTGCCTTTTC CCGCATACCA 601 CGGCGCGGG TCGTCGGCA GTCCGTCGAT AAAGGCAAGG TTATTGCCTT 651 CGCCCTGCAC ATCGGGAACA TTCCCCCAAA AATCATAGCC GTCATCGGGC 701 AACTCGTCGG TTTCGATACC CGTCCAACTG CCGAATCCGC GTAA

This corresponds to the amino acid sequence <SEQ ID 256; ORF 082>: m082.pep

- 1 MXLLKLPAVA NTASSPKRRR NTAASISFTV VLPPEPVMPN TNGFTFSRHA
- 51 FASVCNAASV SSTFNAPSIA AOSSRETTTA AAPAANTSST KSCASNRSPA
- 101 NAKNTSPARM SRLSVTMRDT GLLSDGIGSL RAWOMKFRSS GFIFTFVNIR
- 151 AADTSVAADF FIACFAVVKH RLFSHSHSXF FLYVSFFRRI FSRFAFSRIP
- 201 RRGVVGQSVD KGKVIAFALH IGNIPPKIIA VIGQLVGFDT RPTAESA*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae

ORF 082 shows 92.7% identity over a 247 aa overlap with a predicted ORF (ORF 082.ng) from N. gonorrhoeae:

m082/g082

m082.pep g082	10 MXLLKLPAVANTASS                   :         MWLLKLPAVAETASS			50 TFSRHAFASV  :     :  TLSRHAFANV 50	
m082.pep	70 SSTFNAPSIAAQSSF          SSTFNAPPKAAQSSF 70	11 11111		111:1:111	
m082.pep	130 GLLSDGIGSLRAWQN   :         GLFSDGIGSLRAWQN 130	:	1111111111	170 FAVVKHRLFS          FAVVKHRLFS 170	11111
m082.pep g082	190 FLYVSFFRRIFSRF/          FLYVSFFRRIFSRF/ 190	[] [][[][]	1111 111:1	111111111	111111
m082.pep	RPTAESAX          RPTAESAX				

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 257>: a082.seq

- ATGTGGTTGT TGAAGTTGCC TGCCGTCGCC AAAACGGCAT TATCGCCGAA
- 51 ACGGCGGCGC AATACCGCAG CCAACATTTC CTTCACCGTC GTCTTGCCGC

BNSDOCID: <WO___9957280A2_|_>

```
CCGAGCCGGT AATACCGAAC ACAAACGGGT TCACATTCTC GCGCCACGCC
          TTCGCCAACA TTTGCAACGC GGTAAGCGTG TCATCGACTT TCAACGCGCC
     151
     201
          ATCCATTGCA ACGCAATCTT CGCGCGAAAC CACAACCGCC GCCGCACCCG
          CAGCCAATAC GTCTTCAACA AAATCATGCG CATCAAACCG CCCGCCCGCC
     251
         AATGCGAAAA ACACATCGCC CGCGCGGATG TCGCGGCTGT CGGTTACGAT
     301
          GCGCGACACG GGTTTGCTTT CAGACGGCAT CGGAAGCTTG AGGGCTTGGC
     351
          AGATGAAATT TAGGTCCAGT GGTTTCATAT TTACTTTCGT TAATATTCGG
     401
          GCGGCGGACA CATCGGTAGC GGCTGATTTT TTTATCGCCT GTTTTGCTGT
     451
          GGTAAAACAC AGATTATTTT CCCATTCTCA TTCGGCATTT TTTCTGTACG
          TATCATTTTT TAGACGTATT TTTAGTCGAT TTGCCTTTTC CCGCATACCA
     551
          CGGCGCGGG TCGTCGGCA GTCCGTCGAT AAAGGCAAGG TTATTGCCTT
     601
          CGCCCTGCAC ATCGGGAACA TTCCCCCAAA AATCATAGCC GTCATCGGGC
     651
     701
          AACTCGTCGG TTTCGATACC CGTCCAACTG CCGAATCCGC GTAA
This corresponds to the amino acid sequence <SEQ ID 258; ORF 082.a>:
a082.pep
         MWLLKLPAVA KTALSPKRRR NTAANISFTV VLPPEPVIPN TNGFTFSRHA
       1
         FANICNAVSV SSTFNAPSIA TQSSRETTTA AAPAANTSST KSCASNRPPA
      51
         NAKNTSPARM SRLSVTMRDT GLLSDGIGSL RAWQMKFRSS GFIFTFVNIR
     101
         AADTSVAADF FIACFAVVKH RLFSHSHSAF FLYVSFFRRI FSRFAFSRIP
     151
         RRGVVGQSVD KGKVIAFALH IGNIPPKIIA VIGQLVGFDT RPTAESA*
     201
            95.5% identity over a 247 aa overlap
m082/a082
                    10
                             20
                                      30
                                               40
                                                        50
m082.pep
            MXLLKLPAVANTASSPKRRRNTAASISFTVVLPPEPVMPNTNGFTFSRHAFASVCNAASV
            MWLLKLPAVAKTALSPKRRRNTAANISFTVVLPPEPVIPNTNGFTFSRHAFANICNAVSV
a082
                   10
                             20
                                      30
                                               40
                                                        50
                                                                 60
                   70
                             80
                                      90
                                              100
                                                       110
                                                                120
            SSTFNAPSIAAQSSRETTTAAAPAANTSSTKSCASNRSPANAKNTSPARMSRLSVTMRDT
m082.pep
            SSTFNAPSIATQSSRETTTAAAPAANTSSTKSCASNRPPANAKNTSPARMSRLSVTMRDT
a082
                   70
                             80
                                      90
                                              100
                                                       110
                  130
                            140
                                     150
                                              160
                                                       170
                                                                180
            GLLSDGIGSLRAWQMKFRSSGFIFTFVNIRAADTSVAADFFIACFAVVKHRLFSHSHSXF
m082.pep
            a082
            {\tt GLLSDGIGSLRAWQMKFRSSGFIFTFVNIRAADTSVAADFFIACFAVVKHRLFSHSHSAF}
                  130
                           140
                                     150
                                              160
                                                       170
                                                                180
                  190
                           200
                                     210
                                              220
                                                       230
                                                                240
            {\tt FLYVSFFRRIFSRFAFSRIPRRGVVGQSVDKGKVIAFALHIGNIPPKIIAVIGQLVGFDT}
m082.pep
            a082
            FLYVSFFRRIFSRFAFSRIPRRGVVGQSVDKGKVIAFALHIGNIPPKIIAVIGQLVGFDT
                  190
                           200
                                    210
                                              220
m082.pep
            RPTAESAX
            a082
           RPTAESAX
```

### The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 259>:

1	ATGAAacaAT	CCGcccgaat	aAAAAATATG	GATCAGACAT	TAAAAAATAc
51	attgggcatt	tGCGCGcttt	tagcctTTTG	TTTTggcgcG	GCCATCGCAT
101	CAGGTTATCA	CTTGGAATAT	GAATACGGCT	ACCGTTATTC	TGCCGTGGGC
151	GCTTTGGCTT	CGGTTGTATT	TTTATTATTA	TTGGCACGCG	GCTTCCCCCC
201	CGTTTCTTCA	GTTGTTTTAC	TGATTTACGT	CGGCACAACC	GCCCTATATT
251	TGCCGGTCGG	CTGGCTGTAT	GGTGCGCCTT	CTTATCAGAT	AGTCGGTTCG
301	ATATTGGAAA	GCAATCCTGC	CGAGGCGCGT	GAATTTGTCG	GCAATCTTCC
351	CGGGTCGCTT	TATTTTGTGC	AGGCATTATT	TTTCATTTTT	GGCTTGACAG

Ţ.

```
TTTGGAAATA TTGTGTATCT GTGGGGGTAT TTGCTGACGT AAAAAACTAT
         451 AAACGTCGCA GCAAAATATG GCTGACCATA TTATTGACTT TGATTTTGTC
         501 CTGCGCGGTG ATGGAGAAAA TCGccggcga taaAGATTGG CGAGaacctg
         551 atgccggcct gttgttgaat ATTTTcgacc tgtattaCga cttggctttc
         601 cqcqccqqca cAATATGCCG CCAAGCGCGC CCAcattttg gaaqCaqcaa
              aaaaagcgtC AACATGGCAt atccgccaac ttgcgcccaa gTAtaa
This corresponds to the amino acid sequence <SEQ ID 260; ORF 084.ng>:
    g084.pep
              MKQSARIKNM DQTLKNTLGI CALLAFCFGA AIASGYHLEY EYGYRYSAVG
              ALASVVFLLL LARGFPRVSS VVLLIYVGTT ALYLPVGWLY GAPSYQIVGS
          51
              ILESNPAEAR EFVGNLPGSL YFVQALFFIF GLTVWKYCVS VGVFADVKNY
         101
             KRRSKIWLTI LLTLILSCAV MEKIAGDKDW REPDAGLLLN IFDLYYDLAF
         151
         201 RAGTICRQAR PHFGSSKKSV NMAYPPTCAQ V*
The following partial DNA sequence was identified in N. meningitidis <SEO ID 261>:
    m084.seq
              ATGAAACAAT CCGCCCGAAT AAAa.ATATG AATCAGACAT TACTTTATAC
           1
              ATTGGGCATT TGCGCGCTTT TAACCTTTnn nnnnnnnnn nnnnnnnnn
          51
             nnnntatca cccngaatat gaatacggct accgttattc tgccgtgggt
         101
             GCTTTGGCTT CGGTTGTATT TTTATTATTA TTGGCACGCG GTTTCCCGCG
         151
         201 CGTTTCTTCA GTTGTTTTAC TGATTTACGT CGGCACAACC GCCCTATATT
              TGCCGGTCGG CTGGCTGTAT GGTGCGCCGT CTTATCAGAT AGTCGGTTCG
         251
         301 ATATTGGAAA GCAATCCTGC CGAGGCGCGT GAATTTGTCG GCAATCTTCC
         351 CGGGTCGCTT TATTTTGTGC AGGCATTATT TTTCATTTTT GGCTTGACAG
         401 TTTGGAAATA TTGTGTATCG GGGGGGGTAT TTGCTGACGT AAAAAACTAT
         451 AAACGCCGCA GCAAAATATG GCTGACTATA TTATTGACTT TGATTTTGTC
         501 CTGCGCGGTG ATGGATAAAA TCGCCAGCGA TAAAGATTTG CGAGAACCTG
         551 ATGCCGGCCT GTTGTTGAAT ATTTTCGACC TGTATTACGA TTTGGCT.TC
             CGCGCCGGCA CAATATGCCG CCAAGCGCGC CCACATTTTG GAAGCAGCAA
              AAAAAGCGTC AACATGGCAT ATCCGTCATG TTGCGCCCAA GTATAA
This corresponds to the amino acid sequence <SEQ ID 262; ORF 084>:
    m084.pep
              MKQSARIKXM NQTLLYTLGI CALLTFXXXX XXXXXYHPEY EYGYRYSAVG
           1
              ALASVVFLLL LARGFPRVSS VVLLIYVGTT ALYLPVGWLY GAPSYQIVGS
          51
              ILESNPAEAR EFVGNLPGSL YFVQALFFIF GLTVWKYCVS GGVFADVKNY
         101
              KRRSKIWLTI LLTLILSCAV MDKIASDKDL REPDAGLLLN IFDLYYDLAX
              RAGTICROAR PHFGSSKKSV NMAYPSCCAO V*
Computer analysis of this amino acid sequence gave the following results:
Homology with a predicted ORF from N. gonorrhoeae
ORF 084 shows 90.5% identity over a 231 aa overlap with a predicted ORF (ORF 084.ng)
from N. gonorrhoeae:
    m084/g084
                                 20
                                                   30
                 MKOSARIKXMNQTLLYTLGICALLTF-----YHPEYEYGYRYSAVGALASVVFLLL
    m084.pep
                 g084
                 MKQSARIKNMDQTLKNTLGICALLAFCFGAAIASGYHLEYEYGYRYSAVGALASVVFILL
                        10
                                 20
                                           30
                                                             50
                                                                       60
                       60
                                 70
                                          80
                                                   90
                                                           100
                 m084.pep
                 g084
                 LARGFPRVSSVVLLIYVGTTALYLPVGWLYGAPSYQIVGSILESNPAEAREFVGNLPGSL
                        70
                                 80
                                           90
                                                   100
                                                            110
                                                                      120
                      120
                                130
                                         140
                                                  150
                                                                     170
                 YFVQALFFIFGLTVWKYCVSGGVFADVKNYKRRSKIWLTILLTLILSCAVMDKIASDKDL
    m084.pep
                 YFVQALFFIFGLTVWKYCVSVGVFADVKNYKRRSKIWLTILLTLILSCAVMEKIAGDKDW
    g084
```

130

140

150

160

170

180

BNSDOCID: <WO___8957280A2_I_>

	180	190	200	210	220
m084.pep	REPDAGLLLNIFI	DLYYDLAXRA	GTICRQARPHF(	SSKKSVNMA	YPSCCAQVX
•••					11 11111
g084	REPDAGLLLNIFI	LYYDLAFRA	GTI CRQARPHFO	SSKKSVNMA	YPPTCAQVX
	190	200	210	220	230

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 263>: a084.seq

1	ATGAAACAAT	CCGCCCGAAT	AAAAAATATG	GATCAGACAT	TAAAAAATAC
51	ATTGGGCATT	TGCGCGCTTT	TAGCCTTTTG		GCCATCGCAT
101	CAGGTTATCA	CTTGGAATAT	GAATACGGCT		TGCCGTGGGT
151	GCTTTGGCTT	CGGTTGTATT	TTTATTATTA		GTTTCCCGCG
201	CGTTTCTTCA	GTTGTTTTAC	TGATTTACGT		GCCCTATATT
251	TGCCGGTCGG	CTGGCTGTAT	GGTGCGCCGT	CTTATCAGAT	
301	ATATTGGAAA	00121100100	CGAGGCGCGT	GAATTTGTCG	GCAATCTTCC
351	CGGGTCGCTT	TATTTTGTGC	AGGCATTATT	TTTCATTTTT	GGCTTGACAG
401	TTTGGAGATA		GGGGGGGTAT	TTGCTGACGT	AAAAAACTAT
451	AAACGCCGCA		GCTGACTATA	TTATTGACTT	TGATTTTGTC
501	CTGCGCGGTG	ATGGATAAAA	TCGCCAGCGA	TAAAGATTTG	CGAGAACCTG
551	ATGCCGGCCT	GTTGTTGAAT	ATTTTCGACC	TGTATTACGA	TTTGGCTTCC
601			CCAAGCGCGC	CCACATTTTG	GAAGCAGCAA
651	AAAAAGCGTC	AACATGGCAT	ATCCGTCATG	TTGCGCCCAA	GTATAA

# This corresponds to the amino acid sequence <SEQ ID 264; ORF 084.a>: a084.pep

1	MKQSARIKNM	DOTLKNTLGI	CALLAFCFGA	AIASGYHLEY	EYGYRYSAVG
51	ALASVVFLLL	LARGFPRVSS	VVLLIYVGTT	ALYLPVGWLY	GAPSYOTVGS
101	ILESNPAEAR	EFVGNLPGSL	YFVQALFFIF	GLTVWRYCVS	GGVFADVKNY
151	KRRSKIWLTI	LLTLILSCAV	MDKIASDKDL	REPDAGLLLN	I FDLYYDLAS

201 XAGTICRQAR PHFGSSKKSV NMAYPSCCAQ V*

### m084/a084 92.2% identity over a 231 aa overlap

m084.pep	10 MKQSARIKXMNQTL	20 LYTLGICALI	30 TFXXXXXXX	40 XYHPEYEYG	50 YRYSAVGAT.AS	60
a084		 KNTLGICALI	:  .AFCFGAAIAS	11 11111	11111111111	111111
	10	20	30	40	50	60
m084.pep	70	80	90	100	110	120
oo 1.pcp	LARGFPRVSSVVLL	LIVGTTALYL	PVGWLYGAPS	SYQIVGSILES	SNPAEAREFVG	NLPGSL
a084			111111111	41111111	1111111111	111111
a004	LARGFPRVSSVVLL:	IYVGTTALYL	PVGWLYGAPS	YQIVGSILES	SNPAEAREFVG	NLPGSL
	70	80	90	100	110	120
	130	140	150	160	170	180
m084.pep	YFVOALFFIFGLTV	KYCVSGGVF	ADVKNYKRRS	KIWLTILLTI	TISCAVMORT	ASDKDI
a084	_	:		1111111111	111111111	111111
	YFVQALFFIFGLTVV	140	ADVENIERS			ASDKDL
	130	140	150	160	170	180
	190	200	210	220	230	
m084.pep	REPDAGLLLNIFDLY	YDLAXRAGT	ICRQARPHFG	SSKKSVNMAY	PSCCAOVX	
			111111111	1111111111	1111111	
a084	REPDAGLLLNIFDLY	YDLASXAGT	ICROARPHFG.	SSKKSVNMAY	PSCCAOVY	
	190	200	210	220	230	

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 265>:

1 ATGGGCAAAG GGCAGGACTT CACGCCCCTG CGCGACGCGT TGAAAGATAA

```
51 GGCAAAAGGC GTGTTCCTGA TCGGCGTCGA TGCGCCGCAA ATCCGCCGCG
          101 ATTTGGACGG CTGCGGCTTG AACCTGACCG ACTGCGTCAC TTTGGAAGAG
               GCGGTTCAGA CGGCATACGC CCAAGCCGAA GCGGGCGATA TTGTCTTGCT
               CAGCCCCGCC TGCGCGAGTT TCGATATGTT TAAAGGCTAC GCGCACCGTT
          201
               CGGAAGTGTT tatCGAAGCG TTTAAGGCTT TGTGA
          251
This corresponds to the amino acid sequence <SEQ ID 266; ORF 085.ng>:
     g085.pep
               MGKGQDFTPL RDALKDKAKG VFLIGVDAPQ IRRDLDGCGL NLTDCVTLEE
            ٦
              AVQTAYAQAE AGDIVLLSPA CASFDMFKGY AHRSEVFIEA FKAL*
           51
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 267>:
     m085.seq
               ATGGGTAAAG GGCAGGACTT CACGCCCCTG CGCGATGCAC TGGTAGGCAA
            1
           51
               GGCAAAAGGC GTGTTCTTGA TTGGTGTCGA TGCGCCGCAA ATCCGCCGCG
          101 ATTTGGACGG CTGCGGCTTG AATATGACCG ACTGCGCCAC TTTGGGAGAA
          151 GCCGTTCAGA CGGCATATGC CCAAGCCGAA GCAGGCGATA TTGTGTTGCT
          201 CAGCCCCGCC TGCGCGAGCT TTGATATGTT CAAAGGCTAC GCGCACCGTT
          251 CGGAAGTGTT TATCGAAGCG TTTAAGGCTT TGTGA
This corresponds to the amino acid sequence <SEQ ID 268; ORF 085>:
     m085.pep
               MGKGQDFTPL RDALVGKAKG VFLIGVDAPQ IRRDLDGCGL NMTDCATLGE
              AVQTAYAQAE AGDIVLLSPA CASFDMFKGY AHRSEVFIEA FKAL*
Computer analysis of this amino acid sequence gave the following results:
Homology with a predicted ORF from N. gonorrhoeae
ORF 085 shows 94.7% identity over a 94 aa overlap with a predicted ORF (ORF 085.ng)
from N. gonorrhoeae:
     m085/g085
                          10
                                    20
                                              30
                                                       40
                                                                 50
                                                                           60
                  MGKGQDFTPLRDALVGKAKGVFLIGVDAPQIRRDLDGCGLNMTDCATLGEAVOTAYAOAE
     m085.pep
                  g085
                  {\tt MGKGQDFTPLRDALKDKAKGVFLIGVDAPQIRRDLDGCGLNLTDCVTLEEAVQTAYAQAE}
                          10
                                    20
                                             30
                                                       40
                                                                 50
                                                                           60
                          70
                                    80
                  AGDIVLLSPACASFDMFKGYAHRSEVFIEAFKALX
     m085.pep
                  g085
                  AGDIVLLSPACASFDMFKGYAHRSEVFIEAFKALX
                          70
                                    80
                                             90
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 269>:
a085.seq
         ATGGGCAAAG GGCAGGACTT CACGCCCCTG CGCGACGCGC TTGCCGGCAA
      1
         GGCAAAAGGC GTGTTCCTGA TCGGTGTCGA TGCGCCGCAA ATCCGCCGCG
         ATTTGGACGG CTGCGATCTG AATATGACCG ACTGCGCCAC TTTGGAAGAA
     101
     151 GCGGTTCAGA AGGCATATGC CCAAGCCGAA GCGGGCGATA TCGTGCTGCT
     201 CAGCCCCGCC TGCGCGAGTT TCGATATGTT TAAAGGCTAC GCGCACCGTT
     251 CGGAAGTGTT TATCGGGGCG TTTAAGGCTT TGTGA
This corresponds to the amino acid sequence <SEQ ID 270; ORF 085.a>:
a085.pep
         MGKGQDFTPL RDALAGKAKG VFLIGVDAPQ IRRDLDGCDL NMTDCATLEE
         AVQKAYAQAE AGDIVLLSPA CASFDMFKGY AHRSEVFIGA FKAL*
             94.7% identity over a 94 aa overlap
m085/a085
                               20
                                         30
                                                  40
            {\tt MGKGQDFTPLRDALVGKAKGVFLIGVDAPQIRRDLDGCGLNMTDCATLGEAVQTAYAQAE}
m085.pep
```

```
MGKGQDFTPLRDALAGKAKGVFLIGVDAPQIRRDLDGCDLNMTDCATLEEAVQKAYAQAE
 a085
                                         30
                      70
                               80
 m085.pep
              AGDIVLLSPACASFDMFKGYAHRSEVFIEAFKALX
              a085
              AGDIVLLSPACASFDMFKGYAHRSEVFIGAFKALX
                     70
                               80
 The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 271>:
      g086.seq
               ATGGTGGTGC TGATGACGGC GTTCGGCCTG CTGATGATTT ATTCGGCTTC
             1
               TGTGTATTTG GCATCGAAGG AAGGCGGCGA TCAGTTTTTC TATTTGACCA
            51
               GGCAGGCGGG GTTCGTCGTT GCCGGCCTTA TAGCGAGCGG TTTTTTATGG
           151 TTTCTTTGCA GGATGAGGAC ATGGCGGCGG CTTGTGCCGT GGATTTTTGC
           201 CTTATCCGGC CTGTTGCTGG TAGCCGTATT GATTGCCGGG CGCGAAATCA
               ATGGCGCGAC CCGTTGGATA CCTTTGGGTC CGTTGAATTT CCAGCCGACC
               GAGCTGTTCA AGCTGGCAGT CATCCTTTAT TTGGCAAGCC TGTTCACGCG
               CCGTGAAGAA GTGTTGCGCA GCATGGAAAG TTTGGGTTGG CAGTCGATTT
               GGCGGGGGAC GGCCAACCTG ATTATGTCCG CCACCAATCC GCAGGCACGT
              CGTGAAACAT TAGAAATGTA CGGCCGTTTC CGGGCGATCA TCCTGCCGAT
           451
          501 TATGCTGGTG GCGTTCGGTT TGGTGCTGAT AATGGTACAG CCGGATTTCG
          551 GTTCGTTTGT CGTCATTACC GTCATTACCG TTGGAATGCT GTTTCTGGCA
          601 GGATTGCCGT GGAAATATTT TTTTGTCCTG GTAGGCAGCG TCTTGGGTGG
          651 GATGGTGCTG ATGATTACCG CCGCTCCCTA CCGTGTGCAG CGGGTAGTGG
          701 CATTTTGGA CCCGTGGAAA GACCCGCAGG GTGCCGGCTA CCAGCTTACC
          751 CACTCTCTGA TGGCAATCGG GCGCGGAGAG TGGTTCGGTA TGGGTTTGGG
          801 TGCGAGTTTG AGCAAACGCG GCTTTCTGCC GGAAGCGCAT ACCGATTTTA
               TTTTTGCCAT CATCGCTGAA GAATTCGGCT TCTTCGGGAT GTGCGTGCTG
               ATATTCTGTT ACGGCTGGCT GGTGGTGCGG GCGTTTTCCA TCGGCAAGCA
              GTCGCGCGAT TTGGGtttgA CTTTCAACGC CTATATCGCT TCGGGTATCG
         1001 GCATTTGGAT CGGTATCCAA AGTTTCTTCA ATATCGGTGT GAACATCGGT
         1051 GCTTTGCCGA CCAAAGGTCT GACGCtgCcg tTGATGTCCT ATGGcggTTC
         1101 GTCAGTCTTT TTCATGCTGA TCAGCATGAT GCTGCTGTTG CGTATCGATT
         1151 ATGAAAACCG CCAGAAAATG CGCGGTTACC GGGTGGAGTA AA
This corresponds to the amino acid sequence <SEQ ID 272; ORF 086.ng>:
     g086.pep
               MVVLMTAFGL LMIYSASVYL ASKEGGDQFF YLTRQAGFVV AGLIASGFLW
               FLCRMRTWRR LVPWIFALSG LLLVAVLIAG REINGATRWI PLGPLNFQPT
           51
              ELFKLAVILY LASLFTRREE VLRSMESLGW QSIWRGTANL IMSATNPQAR
          151 RETLEMYGRF RAIILPIMLV AFGLVLIMVQ PDFGSFVVIT VITVGMLFLA
              GLPWKYFFVL VGSVLGGMVL MITAAPYRVQ RVVAFLDPWK DPQGAGYQLT
          251 HSLMAIGRGE WFGMGLGASL SKRGFLPEAH TDFIFAIIAE EFGFFGMCVL
          301 IFCYGWLVVR AFSIGKQSRD LGLTFNAYIA SGIGIWIGIQ SF<u>FNIGVNIG</u>
              ALPTKGLTLP LMSYGGSSVF FMLISMMLLL RIDYENRQKM RGYRVE*
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 273>:
     m086.seq
              ATGGTGGTGC TGATGACGGC GTTCAGCCTG CTGATGATTT ATTCGGCTTC
              TGTGTATTTG GCATCAAAAG AAGGCGGCGA TCAGTTTTTC TATTTGACCA
           51
          101 GACAGGCGGG GTTCGTCGTT GCCGGCTTGA TAGCGAGCGG TTTGTTATGG
              TTTCTTTGCA GGATGAGGAC ATGGCGGCGG CTTGTGCCGT GGATTTTTGC
          151
              CCTATCCGGC CTGTTGCTGG TAGTCGTATT GATTGCCGGG CGCGAAATCA
              ATGGCGCGAC CCGTTGGATA CCTTTGGGTC CGTTGAATTT CCAGCCGACC
          251
              GAGCTGTTCA AGCLGGCGGT CATCCTTTAT TTGGCAAGCC TGTTCACGCG
         301
         351
             CCGTGAAGAA GTGTTGCGCA GCATGGAAAG TTTGGGTTGG CAGTCGATTT
         401 GGCGGGGGAC GGCCAATCTG ATCATGTCCG CCACCAATCC GCAGrCACGT
         451 CGTGAAACAT TAGAAATGTA CGGCCGTwTC CGGGCGATCA TCCTGCCGAT
         501 TATGCTGGTG GCGTTCGGTT TGGTGCTGAT AATGGTACAG CCGGATTTCG
```

551	GTTCGTTTGT	CGTCATTACC	GTCATTGCCG	TTGGAATGCT	GTTTTTGGCA
601	GGATTGCCGT	GGAAATATTT	TTTCGTCCTG	GTAGGCAGCG	TCTTGGGCGG
651	GATGGTGCTG	ATGATTACCG	CCGCTCCCTA	CCGTGTGCAG	CGGGTAGTGG
701	CATTTTTGGA	CCCGTGGAAA	GACCCGCAGG	GTGCCGGCTA	CCAGCTTACC
751	CACTCTCTGA	TGGCAATCGG	GCGCGGAGAG	TGGTTCGGTA	TGGGTTTGGG
801	TGCGAGTTTG	AGCAAACGCG	GCTTTCTGCC	GGAAGCGCAT	ACCGATTTTA
851	TTTTTGCCAT	CATCGCCGAA	GAATTCGGTT	TCTTCGGTAT	GTGCGTGCTG
901	ATATTCTGTT	ACGGCTGGCT	GGTGGTGCGG	GCGTTTTCCA	TCGGCAAGCA
951	GTCGCGCGAT	TTGGGTTTGA	CTTTCAACGC	CTATATCGCT	TCGGGTATCG
1001	GCATTTGGAT	CGGKTTCCAA	AGTTTCTTCA	ATATCGGTGT	GAACATCGGT
1051	GCTTTGCCGA	mCAAAgGyCT	GACGCyGCCG	Tg.AtGTCCw	ATGGCGGTTC
1101	GTCAGTCTTT	TTCATGCTGA	TCAGCATGAT	GCTGCTGTkG	CGTATAGATT
1151	ATGAAAACCG	CCGGAAAATG	CGCGGTTATC	GGGTGGAGTA	A

This corresponds to the amino acid sequence <SEQ ID 274; ORF 086>: m086.pep

1 MVVLMTAFSL LMIYSASVYL ASKEGGDQFF YLTRQAGFVV AGLIASGLLW
51 FLCRMRTWRR LVPWIFALSG LLLVVVLIAG REINGATRWI PLGPLNFQPT
101 ELFKLAVILY LASLFTRREE VLRSMESLGW QSIWRGTANL IMSATNPQXR
151 RETLEMYGRX RAIILPIMLV AFGLVLIMVQ PDFGSFVVIT VIAVGMLFLA
201 GLPWKYFFVL VGSVLGGMVL MITAAPYRVQ RVVAFLDPWK DPQGAGYQLT
251 HSLMAIGRGE WFGMGLGASL SKRGFLPEAH TDFIFAIIAE EFGFFGMCVL
301 IFCYGWLVVR AFSIGKQSRD LGLTFNAYIA SGIGIWIGXQ SFFNIGVNIG
351 ALPXKGLTXP XMSXGGSSVF FMLISMMLLX RIDYENRRKM RGYRVE*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae

10

ORF 086 shows 96.7% identity over a 396 aa overlap with a predicted ORF (ORF 086.ng) from N. gonorrhoeae:

30

40

50

60

20

m086/g086

	10	20	30	40	50	60
m086.pep	MVVLMTAFSLLMIYS	ASVYLASKE	GDQFFYLTR	QAGFVVAGLI.	ASGLLWFLCR	MRTWRR
					111:11111	111111
g086	MVVLMTAFGLLMIYS	ASVYLASKE	GDQFFYLTR	QAGFVVAGLI.	ASGFLWFLCR	MRTWRR
	10	20	30	40	50	60
	70	80	90	100	110	120
m086.pep	LVPWI FALSGLLLVV	VLIAGREING	ATRWIPLGP:	LNFQPTELFK	LAVILYLASL	FTRREE
	111111111111:	11111111			1111111111	111111
g086	LVPWIFALSGLLLVA	VLIAGREIN	SATRWIPLGP:	LNFOPTELFK	LAVILYLASI	FTRREE
5	70	80	90	100	110	120
	130	140	150	160	170	180
m086.pep	VLRSMESLGWQSIWR	GTANLIMSA:	INPOXRRETL	EMYGRXRAII:	LPIMLVAFGL	VLIMVO
* *		11111111	111Î 1111E	11111 1111	111111111	11111
g086	VLRSMESLGWOSIWR	GTANLIMSA	INPOARRETL	EMYGRFRAII:	LPIMLVAFGL	
2	130	140	150	160	170	180
					_,_	
	190	200	210	220	230	240
m086.pep	PDFGSFVVITVIAVG					
g086	PDFGSFVVITVITVG	MLFLAGLPW	KYFFVIJVGSV	I.GGMVI.MTTA		
<b>3</b>	190	200	210	220	230	240
					250	240
	250	260	270	280	290	300
m086.pep	DPQGAGYQLTHSLMA					
					1111111111	
g086	DPQGAGYQLTHSLMA	TGRGEWEGM	T.GAST.SKRG	FLOFAHTOFT	PATTARERCE	
3	250	260	270	280	290	300
	230	200	2.0	200	230	300
	310	320	330	340	350	360
	J. V	220	330	340	330	200

m086.pep	IFCYGWLVVRAFS]	GKQSRDLGL	TFNAYIASGI	GIWIGXQSFFN	NIGVNIGALPX	KKGLTXP
g086				<u> </u>		. [ [ ] ] .
	310	320	330	340	350	360
	370	380	390			
m086.pep	XMSXGGSSVFFMLI					
G096	I MOVOGO GIVETTUS	1111		111		
g086 LMSYGGSSVFFMLISMMLLLRIDYENROKMRGYRVEX						
	370	380	390			

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 275>: a086.seq

```
ATGGTGGTGC TGATGACGGC GTTCAGCCTG CTGATGATTT ATTCGGCTTC
   1
      TGTGTATTTG GCATCAAAAG AAGGCGGCGA TCAGTTTTTC TATTTGACCA
   51
      GACAGGCGGG GTTCGTCGTT GCCGGCTTGA TAGCGAGCGG TTTGTTATGG
 101
 151 TTTCTTTGCA GGATGAGGAC ATGGCGGCGG CTTGTGCCGT GGATTTTTGC
 201 CCTATCCGGC CTGTTGCTGG TAGTCGTATT GATTGCCGGG CGCGAAATCA
 251 ATGGCGCGAC CCGTTGGATA CCTTTGGGTC CGTTGAATTT CCAGCCGACC
 301 GAGCTGTTCA AGCTGGCGGT CATCCTTTAT TTGGCAAGCC TGTTCACGCG
351 CCGTGAAGAA GTGTTGCGCA GCATGGAAAG TTTGGGTTGG CAGTCGATTT
 401 GGCGGGGGAC GGCCAATCTG ATCATGTCCG CCACCAATCC GCAGGCACGT
 451 CGTGAAACAT TAGAAATGTA CGGCCGTTTC CGGGCGATCA TCCTGCCGAT
 501 TATGCTGGTG GCGTTCGGTT TGGTGCTGAT AATGGTACAG CCGGATTTCG
      GTTCGTTTGT CGTCATTACC GTCATTGCCG TTGGAATGCT GTTTTTGGCA
      GGATTGCCGT GGAAATATTT TTTCGTCCTG GTAGGCAGCG TCTTGGGCGG
 651 GATGGTGCTG ATGATTACCG CCGCTCCCTA CCGTGTGCAG CGGGTAGTGG
     CATTTTTGGA CCCGTGGAAA GACCCGCAGG GTGCCGGCTA CCAGCTTACC
     CACTCTCTGA TGGCAATCGG GCGCGGAGAG TGGTTCGGTA TGGGTTTGGG
      TGCGAGTTTG AGCAAACGCG GCTTTCTGCC GGAAGCGCAT ACCGATTTTA
 851
      TTTTTGCCAT CATCGCCGAA GAATTCGGTT TCTTCGGTAT GTGCGTGCTG
 901 ATATTCTGTT ACGGCTGGCT GGTGGTGCGG GCGTTTTCCA TCGGCAAGCA
 951 GTCGCGCGAT TTGGGTTTGA CTTTCAACGC CTATATCGCT TCGGGTATCG
1001
     GCATTTGGAT CGGTATCCAA AGTTTCTTCA ATATCGGTGT GAACATCGGT
      GCTTTGCCGA CCAAAGGTCT GACGCTGCCG TTGATGTCCT ATGGCGGTTC
      GTCAGTCTTT TTCATGCTGA TCAGCATGAT GCTGCTGTTG CGTATAGATT
1101
1151 ATGAAAACCG CCGGAAAATG CGCGGTTACC GGGTGGAGTA A
```

### This corresponds to the amino acid sequence <SEQ ID 276; ORF 086.a>: a086.pep

_					
1	MVVLMTAFSL	LMIYSASVYL	ASKEGGDQFF	YLTROAGFVV	AGI.TASGI.I.W
51	FLCRMRTWRR	LVPWIFALSG	LLLVVVLIAG	REINGATRWI	PLCPINEODE
101	ELFKLAVILY	LASLETERE	VLRSMESLGW	OCTUDOMENT	THEFTHEOP
1 - 1		E .ODI IIIIE	ATUMESTER	OSIMKGIANL	IMSATNPQAR
151	RETLEMYGRF	RAIILPIMLV	AFGLVLIMVO	PDFGSFVVIT	VTAVCMI.FI.A
201	GLPWKYFFVL	VGSVLGGMVI.	MITAAPYRUO	DIMM ET DOWN	DDOCAGUETA
251	TIGT NO TODOR		1111111111VQ	WAANT PDEMV	DEGCACAOT.
	HSLMAIGRGE	WEGMGLGASL	SKRGFLPEAH	TDFIFAITAE	EFGFFGMCVI
301	<u>IFCYGWLVV</u> R	AFSTCKOSPD	TOTTEMATE	CCICINICIO	BI GIT GITC VE
251		TIT DI ONQUIND	DGLIENALIA	SCICIMICIÓ	SFFNIGVNIG
351	ALPTKGLTLP	LMSYGGSSVF	FMLISMMLLL	RIDYENRRKM	RGYRVF*

#### m086/a086 98.0% identity over a 396 aa overlap

m086.pep	10 MVVLMTAFSLLMI !!!!!!!!!! MVVLMTAFSLLMI		1		  ASGLLWFLCR	
		20	30	40	50	60
m086.pep	70	80	90	100	110	120
шестрер	LVPWIFALSGLLLY	VVLIAGREIN	SATRWIPLGE	PLNFQPTELFK	LAVILYLASL	FTRREE
a086	LVPWIFALSGLLLV	VVLIAGREIN		  LNFQPTELFK		 FTRREE

	70	80	90	100	110	120
m086.pep	130 VLRSMESLGWQSIV	140 RGTANLIMSA	150 ATNPOXRRETI	160 LEMYGRXRAII	170 LPIMLVAFG	180 LVLIMVO
a086	VLRSMESLGWQSIV	RGTANLIMSA	ATNPQARRETI	LEMYGRFRAII	LPIMLVAFG	LVLIMVQ
	130	140	150	160	170	180
	190	200	210	220	230	240
m086.pep	PDFGSFVVITVIAV					
mooo.pep	11111111111111	IIIIIIIII			_	HILLIIII METDEMV
a086	PDFGSFVVITVIAV	GMT.FT.AGT.PW				A FI I I I I I I
2000	190	200	210	220	230	240
	250	200	210	220	250	240
	250	260	. 270	280	290	300
m086.pep	DPQGAGYQLTHSLM	AIGRGEWFGM	IGLGASLSKRO	FLPEAHTDFI	FAIIAEEFG	FFGMCVL
		1111111111	411111111		111111111	
a086	DPQGAGYQLTHSLM				FAIIAEEFG:	FFGMCVL
	250	260	270	280	290	300
	310	320	330	340	350	360
m086.pep	IFCYGWLVVRAFSI					
moso.pep	IIIIIIIIIIIIIIIII	IIIIIIIIIIII	IIIIIIIIII	TMIGVÕSEEN	IGVNIGALP.	ANGLIAP
a086	IFCYGWLVVRAFSI	CKOSBDICIT	IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	IIII IIIII TWTGTOSFFN	TCVNTCNTD	יווון ו דער זיד די
4000	310	320	330	340	350	360
	010	<b>32</b> 0	550	340	330	300
	370	380	390			
m086.pep	XMSXGGSSVFFMLI	SMMLLXRIDY	ENRRKMRGYF	RVEX		
• •	11 11 11 11 11 11	11111 1111	111111111	111		
a086	LMSYGGSSVFFMLI	SMMLLLRIDY	ENRRKMRGYF	RVEX		
	370	380	390			

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 277>: g087.seq

```
ATGGGCGGTA AAACCTTTAT GCTGATGGCG GGCGGAACGG GCGGACACAT
  51 TTTCCCAGCT CTGGCTGTGG CGGATTCATT GCGCGTGCGC GGTCATCATG
 101 TAATTTGGCT GGGCAGCAAG GATTCGATGG AAGAGCGCAT CGTGCCGCAA
 151 TACGGCATAC GCTTGGAAAC GCTGGCGATT AAAGGAATAC GCGGCAACGG
 201 CATCAAACGC AAGCTGATGC TTCCGTTTAC TCTGTACAAA ACCGTCCGCG
 251 AAGCGCAGCG GATTATCCGC AAACACCGTG TCGAGTGCGT CATCGGCTTC
     GGCGGTTTTG TTACCTTTCC CGGCGGTCTG GCGGCGAAAC TCTTGGGCGT
 301
 351
     GCCGATTGTG ATTCACGAGC AAAACGCCGT GGCAGGCTTG TCCAACCGCC
 401 ACCTGTCGCg ctGGGCGAAA CGGGTGTTGT ACGCTTTTCC GAAAGCGTTC
 451 AGCCACGAAG GCGGTTTGGT CGGCAACCCC GTCCGCGCCG ATATTAGCAA
501 CCTGCCCGTG CCTGCCGAAC GCTTCCAAGG GCGCGAAGGC CGTCTGAAAA
551 TTTTGGTGGT CGGCGGCAGT TTGGGTGCGG ACGTTTTGAA CAAAACCGTA
 601 CCGCAGGCGT TGGCACTGCT GCCTGAAGAG GTGCGCCCGC AGATGTACCA
 651 CCAGTCGGGG CGTAACAAGC TGGGCAATCT TCAGGCGGAT TATGACGCGT
 701 TGGGCGTGAA AGCGGAATGC GTGGAATTTA TTACCGACAT GGTGTCCGCC
 751 TACCGTGATG CCGATTTGGT GATTTGCCGT GCCGGCGCGC TGACGATTGC
 801 CGAGTTGACG GCGGCGGGC TGGGCGCGTT GTTAGTGCCG TATCCTCACG
 851 CCGTTGATGA CCATCAAACC GCCAACGCGC GTTTCATGGT GCAGGCAGAA
     GCGGGGCTGC TGTTGCCGCA AACCCAGTTG ACGGCGGAAA AACTCGCCGA
 951 AATCCTCGGC AGCCTCAACC GCGAAAAATG CCTCAAATGG GCGGAAAACG
     CCCGTACGTT GGCATTGCCG CACAGCGCGG ATGACGTTGC CGAAGCCGCG
1051
     ATTGCGTGTG CGGCGTAAA
```

This corresponds to the amino acid sequence <SEQ ID 278; ORF 087.ng>: g087.pep

- 1 MGGKTFMLMA GGTGGHIFPA LAVADSLRVR GHHVIWLGSK DSMEERIVPQ 51 YGIRLETLAI KGIRGNGIKR KLMLPFTLYK TVREAQRIIR KHRVE<u>CVIGF</u> 101 <u>GGFVTFPGGL AA</u>KLLGVPIV IHEQNAVAGL SNRHLSRWAK RVLYAFPKAF
- 151 SHEGGLVGNP VRADISNLPV PAERFOGREG RLKILVVGGS LGADVLNKTV

278

```
201 PQALALLPEE VRPQMYHQSG RNKLGNLQAD YDALGVKAEC VEFITDMVSA
                YRDADLVICR AGALTIAELT AAGLGALLVP YPHAVDDHQT ANARFMVQAE
                AGLLLPOTOL TAEKLAEILG SLNREKCLKW AENARTLALP HSADDVAEAA
           351
                IACAA*
 The following partial DNA sequence was identified in N. meningitidis <SEQ ID 279>:
      m087.seq
               ATGGGCGGTA AAACCTTTAT GCTGAWKKCG GGCGGAACGG GCGGACATAT
               TTTCCCCGCG CTGGCGGTGG CGGATTCATT GCGCGCGCGC GGCCATCATG
            51
               TGATTTGGCT GGGCAGCAAG GATTCGATGG AAGAGCGTAT CGTGCCGCAA
           101
               TACGGCATAC GCTTGGAAAC GCTGGCGATT AAAGGCGTGC GCGGCAACGG
           151
               CATCAAACGC AAACTGATGC TGCCGGTTAC TTTGTATCAA ACCGTCCGCG
           201
           251 AAGCGCAGCG GATTATCCGC AAACACCGTG TCGAGTGCGT CATCGGCTTC
           301 GGCGGCTTCG TTACCTTCCC CGGCGGTTTG GCGGCGAAGC TATTAYGCGT
           351 GCCGATTGTG ATTCACGAGC AAAACGCCGT GGCAGGTTTG TCCAACCGCC
          401 ACCTGTCGCG CTGGGCGAAG CGGGTGTTGT ACGCTTTTCC GAAAGCGTTC
          451 AGCCACGAAG GCGGCTTGGT CGGCAACCCC GTCCGCGCCG ATATTAGCAA
          501 CCTGCCCGTG CCTGCCGAAC GCTTCCAAGG GCGTGAAGGC CGTCTGAAAA
          551 TTTTGGTGGT CGGCGCAGT TTGGGCGCGG ACGTTTTGAA CAAAACCGTA
              CCGCATGCAT TGGCTTTGCT GCCCGACAAT GCGCGTCCGC ATATGTACCA
               CCAATCGGGA CGGGGCAAGC TGGGCATCTT GCAGGCGnnn nnnnnnnnn
               701
               nnngcgggat tgggtgcgtt gttagtgccg tatcctcacg cggttgacga
          751
               TCACCAAACC GCCAACGCGC GTTTTATGGT GCAGGCGGAG GCGGGATTGC
               TGTTGCCGCA AACCCAGTTG ACGGCGGAAA AACTCGCCGA GATTCTCGGC
               GGCTTAAACC GCGAAAAATG CCTCAAATGG GCAGAAAACG CCCGTACGTT
               GGCACTGCCG CACAGTGCGG ACGACGTGGC GGAAGCCGCG ATTGCGTGTG
         1001
               CGGCGTAA
This corresponds to the amino acid sequence <SEQ ID 280; ORF 087>:
     m087.pep
              MGGKTFMLXX GGTGGHIFPA LAVADSLRAR GHHVIWLGSK DSMEERIVPQ
              YGIRLETLAI KGVRGNGIKR KLMLPVTLYQ TVREAQRIIR KHRVECVIGF
           51
              GGFVTFPGGL AAKLLXVPIV IHEQNAVAGL SNRHLSRWAK RVLYAFPKAF
          101
              SHEGGLVGNP VRADISNLPV PAERFQGREG RLKILVVGGS LGADVLNKTV
          151
              PHALALLPDN ARPHMYHQSG RGKLGILQAX XXXXXXXXXX XXXXXXXXX
          201
              XAGLGALLVP YPHAVDDHQT ANARFMVQAE AGLLLPQTQL TAEKLAEILG
              GLNREKCLKW AENARTLALP HSADDVAEAA IACAA*
Computer analysis of this amino acid sequence gave the following results:
Homology with a predicted ORF from N. gonorrhoeae
ORF 087 shows 83.9% identity over a 355 aa overlap with a predicted ORF (ORF 087.ng)
from N. gonorrhoeae:
     m087/q087
                                  20
                                            30
                                                     40
                 MGGKTFMLXXGGTGGHIFPALAVADSLRARGHHVIWLGSKDSMEERIVPQYGIRLETLAI
    m087.pep
                 MGGKTFMLMAGGTGGHIFPALAVADSLRVRGHHVIWLGSKDSMEERIVPQYGIRLETLAI
    9087
                         10
                                           30
                                                     40
                                                              50
                                                                        60
                        70
                                  80
                                           90
                                                    100
                                                             110
                 KGVRGNGIKRKLMLPVTLYQTVREAQRIIRKHRVECVIGFGGFVTFPGGLAAKLLXVPIV
    m087.pep
                 KGIRGNGIKRKLMLPFTLYKTVREAQRIIRKHRVECVIGFGGFVTFPGGLAAKLLGVPIV
    g087
                        70
                                  80
                                           90
                                                   100
                                                             110
                                                                       120
                                 140
                                          150
                                                   160
                IHEQNAVAGLSNRHLSRWAKRVLYAFPKAFSHEGGLVGNPVRADISNLPVPAERFQGREG
```

IHEQNAVAGLSNRHLSRWAKRVLYAFPKAFSHEGGLVGNPVRADISNLPVPAERFQGREG

160

170

150

m087.pep

130

140

g087

190   200   210   220   229							
RESTANCE AGGEGETA AAACCTTAT GCTGATGGC GGCGGAACGG GCGCACATAT AAAGCGCTG CTGGCGGAA GATTATGGCG AAGGCGCTC CTGGCGAAC GATTATCCGC AAACCCCTG CGGCGATTCC TGGCCGCAACG GCGCACCAC CGGCGCTTC CGCGCGAACG GCGCACCGC CGCGCCTC CGCGCCAACG GCGCACCTC CGCGCGAACGC CGCCCCC CGCGCCACC CGCGCCAACG CGCGCACCC CGCCCCC CGCGCCAACCC CGCGCAACCC CGCCCCC CGCCCCC CGCCCCCC CGCCCCCC CGCCCCCC			190	200	210		229
REKILVVGGSLGADVLNRTVPQALALLPEEVRPOMYHQSGRNKLGNLQADYDALGVKAEC 190 200 210 220 230 240  220 230 240  230 240 250  m087.pep	· m	1087.pep					
190   200   210   220   230   240   250   230   240   250   230   240   250   230   240   250   230   240   250   230   240   250   230   240   250   250   250   250   250   250   250   250   250   250   250   300   300   250   250   250   250   250   300   300   250   250   250   250   250   250   300   300   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250   250							
m087.pep	9	087	RLKILVVGGSLGA	DVLNKTVPQAL	ALLPEEVRPQM	YHQSGRNKLGN	LQADYDALGVKAEC
g087 VEFITDMVSAYRDADLVICRAGALTIAELTRAGLGALLVPYPHAVDDHQTANARFMVQAE g087 VEFITDMVSAYRDADLVICRAGALTIAELTRAGLGALLVPYPHAVDDHQTANARFMVQAE z50 260 270 280 290 300 m087.pep AGLLPQTQLTAEKLAETLGGLNREKCLKWAENARTLALPHSADDVAEAATACAAX			190	200	210	220	230 240
g087 VEFITDMVSAYRDADLVICRAGALTIAELTRAGLGALLVPYPHAVDDHOTANARFMVQAE g087 VEFITDMVSAYRDADLVICRAGALTIAELTRAGLGALLVPYPHAVDDHOTANARFMVQAE z50 260 270 280 290 300  260 270 280 290 300 310  m087.pep AGLLPOTQLTAEKLAETIGGLNREKCLKWAENARTLALPHSADDVAEAATACAAX							
### PROPRIES OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTROL OF CONTRO					230	240	250
250 260 270 280 290 300 310  ***Triangle Company**	m	1087.pep					
### Table					1111		
m087.pep  AGLLLPQTQLTAEKLAEILGGLNREKCLKWAENARTLALPHSADDVAEAAIACAAX	g	087	VEFITDMVSAYRD	ADLVICRAGAL	TIAELTAAGLG	ALLVPYPHAVD	DHQTANARFMVQAE
### AGLLLPQTQLTAEKLAEILGGLNREKCLKWAENARTLALPHSADDVAEAAIACAAX			250	260	270	280	290 300
### AGLLLPQTQLTAEKLAEILGGLNREKCLKWAENARTLALPHSADDVAEAAIACAAX							
The following partial DNA sequence was identified in N. meningitidis <seq 281="" id="">:  a087.seq  1 ATGGGCGTA AAACCTTTAT GCTGATGGCG GGCGGAACGG GCGCATCATG 101 TAATTGGCT GGGCAGCAAG GATTCGATG GCGCGCGCG GGCCATCATG 101 TAATTGGCT GGCAGCAAG GATTCGATG AAAGCGCTT 201 CATCAAACGC AAGCTGATGC TGCCGTTTAC TTTCTTATCA ACTCTCCGCG 201 CATCAAACGC AAGTTCCTCC CAACCGGT TGCGCGTCG GCGCAACAGG 201 GGCGGCTTCG TTACCTTTCC CGGCGGTTG GCGGCAGCAT TATTAGCCAC 201 GCCGATTGTG AACACCGTG TCGAGTGCGT CATCGGCTTC 301 GGCGGCTTCG TTACCTTTCC CGGCGGTTTG GCGGCAACGT TATTAGCCAC 401 ACCTGTCGCG CTGGCCGAAC GGTTTCC CGGCGGCAACTG 401 ACCTGTCGCG CTGGCCGAAC GCGGCAACCCC GTCCCGCCC AAACACCCCTG GCAGGTTTC TCCAACCGCC 401 ACCTGTCGCG CTGGCCGAAC GCTTCCAAGG CGGCAACACG CGGCGCAACGT TATTAGCAA 501 CCGCCGCGTG CTGGCCGAAC GCTTCCAAGG CGTTTGAACAACCCTA 501 CCGCCGCGT TGGCCGAAC GCTTCCAAGG CGCTCGAAAACCCTA 601 CCGCAGGCAT TGGCTTTGCT GCCGCACACT TATTAGCAA 651 CCAATCGGA CGGCGCAAC TTGGGCGCG ACGTTTGAACAAACCCTA 651 CCAATCGGA CGGGCAAC TGGCCCACAAT GCGCCTCCCGC AAAACCCTA 651 CCAATCGGA CGGGCAAC TGGGCAACT TACCGATGCGC 701 TGGGCGTGCA ACCGGATGC TGGGCAGATT TACCGATACCGC 851 TACCGCCAAC GCGGCAACT TGGCTCCCCC 651 CCAATCGGA CGGGCAACC TGGGCCACAT GCGCCCCCC AAAACCCTA 652 CCAATCGGA CGGGCAAC TGGGCCACAT GCGCCCCCC AAACCCCTA 653 CCAATCGCCA ACCCGATTGCT GAGCGCGAT TACCACCGC 701 TGGGCGTGCA ACCCGGCGCAT TACCACCCC 701 TGGGCGTGCA ACCCGCGCAACT TACCGCACCC 702 TGGGCGTCA ACCCGCGCAACT TACCACCCCC 703 TACCCCCACAC TGGCCGAAC TGGGCGCAAC TGCAGCACCC 801 CCGCTCCCCC CCGATTTGGT GAGCGCGAAC TTACCTCACG 801 CCGCTCCCCC CCGCCCCC TTACCTCACG CCCACACCCCC TTACCTCACG 801 CCGCTCCCCCC TTTCCCCCCCC TTACCCCCC 801 CCGCCCCCCC TTTCCCCCCCC TTTCCCCCCCCCCCC</seq>			260 <b>27</b> 0	280	290	300	310
The following partial DNA sequence was identified in N. meningitidis <seq 281="" id="">:  a087.seq  1 ATGGGCGTA AAACCTTTAT GCTGATGGCG GGCGGAACGG GCGCATCATG 101 TAATTGGCT GGGCAGCAAG GATTCGATG GCGCGCGCG GGCCATCATG 101 TAATTGGCT GGCAGCAAG GATTCGATG AAAGCGCTT 201 CATCAAACGC AAGCTGATGC TGCCGTTTAC TTTCTTATCA ACTCTCCGCG 201 CATCAAACGC AAGTTCCTCC CAACCGGT TGCGCGTCG GCGCAACAGG 201 GGCGGCTTCG TTACCTTTCC CGGCGGTTG GCGGCAGCAT TATTAGCCAC 201 GCCGATTGTG AACACCGTG TCGAGTGCGT CATCGGCTTC 301 GGCGGCTTCG TTACCTTTCC CGGCGGTTTG GCGGCAACGT TATTAGCCAC 401 ACCTGTCGCG CTGGCCGAAC GGTTTCC CGGCGGCAACTG 401 ACCTGTCGCG CTGGCCGAAC GCGGCAACCCC GTCCCGCCC AAACACCCCTG GCAGGTTTC TCCAACCGCC 401 ACCTGTCGCG CTGGCCGAAC GCTTCCAAGG CGGCAACACG CGGCGCAACGT TATTAGCAA 501 CCGCCGCGTG CTGGCCGAAC GCTTCCAAGG CGTTTGAACAACCCTA 501 CCGCCGCGT TGGCCGAAC GCTTCCAAGG CGCTCGAAAACCCTA 601 CCGCAGGCAT TGGCTTTGCT GCCGCACACT TATTAGCAA 651 CCAATCGGA CGGCGCAAC TTGGGCGCG ACGTTTGAACAAACCCTA 651 CCAATCGGA CGGGCAAC TGGCCCACAAT GCGCCTCCCGC AAAACCCTA 651 CCAATCGGA CGGGCAAC TGGGCAACT TACCGATGCGC 701 TGGGCGTGCA ACCGGATGC TGGGCAGATT TACCGATACCGC 851 TACCGCCAAC GCGGCAACT TGGCTCCCCC 651 CCAATCGGA CGGGCAACC TGGGCCACAT GCGCCCCCC AAAACCCTA 652 CCAATCGGA CGGGCAAC TGGGCCACAT GCGCCCCCC AAACCCCTA 653 CCAATCGCCA ACCCGATTGCT GAGCGCGAT TACCACCGC 701 TGGGCGTGCA ACCCGGCGCAT TACCACCCC 701 TGGGCGTGCA ACCCGCGCAACT TACCGCACCC 702 TGGGCGTCA ACCCGCGCAACT TACCACCCCC 703 TACCCCCACAC TGGCCGAAC TGGGCGCAAC TGCAGCACCC 801 CCGCTCCCCC CCGATTTGGT GAGCGCGAAC TTACCTCACG 801 CCGCTCCCCC CCGCCCCC TTACCTCACG CCCACACCCCC TTACCTCACG 801 CCGCTCCCCCC TTTCCCCCCCC TTACCCCCC 801 CCGCCCCCCC TTTCCCCCCCC TTTCCCCCCCCCCCC</seq>	m	087.pep	AGLLLPOTOLTAE	KLAEILGGLNR	EKCLKWAENAR	TLALPHSADDV.	AEAAIACAAX
The following partial DNA sequence was identified in N. meningitidis <seq 281="" id="">:  a087.seq  1 ATGGCGGTA AAACCTTTAT GCGCGGCGGG GGCGACATAT 51 TTTCCCCGCG CTGGCGGTGG CATCCATGG GATTCATT GCGCGCGCCAA 151 TAATTTGGCT GGGCAACAG GATTCATG GAAGCGCAT CATCCACGGC 201 CATCAAACGC AAGCTGATGC TGCCCGAAC 201 CATCAAACGC AAGCTGATGC TGCCGCACA 301 GGCGGCTTCG TACCCTTCC CGGCGTTG CGGCAACTT TATTAGCCGT 301 GCGGCTTCG TACCCTTCC CGGCGGTTG CGGCAACTT TATTAGCCT 301 GCCGATTCTG ATTCACCGC AAACACCGTG TCGACGCGC CATCCGCCCC 401 ACCTGTCGCG CTGGCGAAC CGGCGGTTTG ACGCTTTCC GAAACCCCC 401 ACCTGTCGCG CTGGCGAAC CGGCGTTTGT ACGCTTTCC GAAACCCCTC 451 AGCCACGAAG GCGGCTTGGT CGGCAACCCC GTCCGCCCC ATATTAGCAA 551 TTTTGGTGGT CGGCGAAC CGTCCAACC GCTCCAACACCCC 451 CCAACCGCC CTGCCGAAC CGGCGCACG ATATTAGCAA 551 TTTTGGTGGT CGGCGAAC CGGCGCAACT TGGCGCAACCCC GTCCAACCCC 751 TACCCCGTG CCGGACC TGGCCAAC CGCCCAACCCC TCCAACCCCC 751 TACCGCGTG CCGGCACC TGGCCAAC TGCGCTTCAACACCCC 751 TACCGCGTG CCGGCGCACT TGGCCACCC TCCAACCCCC 751 TACCGCGTG CCGGCGCACT TGGCCGCCG ACGTTTTGAA CAAAACCCTA 661 CCAATCGGGA CGGGCAACC TGGCCAACT GCCGCGCCC AGATTACCA 662 CCAATCGGGA CCGGCGCACT TGGCCGCCC ACGTTTTGAA CAAAACCCTA 663 CCAATCGGGA CGGGCAACC TGGCCGCCC TCCACCCC 751 TACCGCGAC CCGGCCGCC TTGGCCGCC TCCCACCC 751 TACCGCGAC CCGGCCGCC TTGGCCGCC TCCCACCC 751 TACCGCGAC CCGACC TGGCCGACC TACCCACCCC 751 TACCGCGAT CCGCCCC CCACCC TCCCGCCCC TGCCGCCCC 751 TACCGCGAT CCGCCGCCC TTGCCCCC 751 TACCGCGAC CCGGCCGCC TTTCCCCCC 751 TACCGCGAC CCGCCCC TCCCACCC TTTCCCGCC 751 TACCGCGAC CCGCCCCC TCCCACCC TCCCCCC 751 TACCGCGAC CCGCCCC TCCCACCC TCCCCCCC 751 TACCGCGAC CCGCCCCC TCCCACCC CCCCCCCCCCC</seq>							
The following partial DNA sequence was identified in N. meningitidis <seq 281="" id="">:  a087.seq  1 ATGGGCGGTA AAACCTTTAT GCTGATGGCG GGCGGACGG GCCGACATAT 51 TTTCCCCGCG CTGGCGGTGG CGGATCATT GCGCGCGCGC GGCCATCATG 101 TAATTTGGCT GGGCACAAG GATTCGATGG AAAGCGCAT CGTGCCGCAA 151 TACGACATCC TGCTCGAAAC GCTGGCGATT AAAGCGCTGC CGGGCAACGG 201 CATCAAACGC AAGCTGATGC TGCCGTTTAC TTTGTATCAA ACTGTCCGCG 251 AAGCGCAGCA GATTATCCGC AAACACCGTG TCGAGTTGCT CATCGGCTTC 301 GGCGGCTTCG TTACCTTTCC CGGCGGTTTG GCGGCGAACT TATTAGGCGT 351 GCCGATTGGT AATCACAGC AAAACACCGTG GCGAGATTT TCCAACCGCC 401 ACCTGTCGCC CTGGCGGAAC GGGTTTTG TCCAACCGCC 451 AGCCACGAAG GCGGCTTGGT CGGCGACCCC GTCCGCGCC ATATTAGCAA 551 CTTTGGTGGT CGGCCGGCAGT TGGGGCGGA ACGTTTTCC GAAAGCGTTC 451 AGCCACGAAG GCGCCAGT TGGGGCCGG ACGTTTTGCA CAAAACCGTA 551 TTTGGGTGT CGGCCGGCAGT TGGGGCCGG ACGTTTTGAA CAAAACCGTA 651 CCAATCGGGA TGGCTTTGCT GCCCGACAT GCGCGCGCC GTCTCAAAA 651 CCAATCGGGA CGGGGCAAGC TGGGCAGCTT GCAGGGGGAT TACACCGCC 751 TACCGCGAT GCGCGGCAGT TGGCCGCGC ACGTTTGAA CAAAACCGTA 651 CCAATCGGGA CCGGAATGC GCGAACTT GCAGGCGGAT TACACCGCC 751 TACCGCGAC CCGGAATGC GTGGAATTA TTACCGATAT GCGCTCGCC 851 CCGATGAC CCGATTTGGT GATTTTCCCGT GCCGGCGCC TACCGCCCC 851 CCGTTGATGA CCATCAAACC GCCAACGCC TTTAGTGC GTACCGCGAA 951 GATTTCCGC GGCTTAAACC GCCAACGCC GTTTTAGTG CAAGCGGAA 951 GATTTCCGC GGCTTAAACC GCCAACGCC GTTTTATGGT GCAGCCGGA 951 GATTTCCGC GGCTTAAACC GCCAACGCC GTTTTATGGT GCAGCCGGA 951 GATTTCCGC GGCTTAAACC GCCAACGCC GTTTTATGGT GCAGCCGGA 951 GATTTCCGC GGCCTAAACC GCCAACGCC GTTTTATGGT GCAGCCGGA 951 GATTTCCGC GGCCTAAACC GCCAACTGC GCAACAGCCC GTTTTATGGT GCAGCCGGAC 1051 ATTGCGTGT GGCACTGCC CACAGTTGC CCCACAGTTGC CGAACCGCC GTTTTATGGT GCAGCCGGAC 1051 ATTGCGTGT GGCACTGCC CACAGTTGC CCACAGTTGC CGAACCGCC GTTTTATGGT GCAGCCGGAC 1051 ATTGCGTGT GGCACTGCC CACAGTTGC CCACAGTTGC CGAACCGCC GTTTATGGT GCAGCCGGAC 1051 ATTGCGTGT GGCACTGCC CACAGTTGC CACCAGTTGC CAGCCGGAAAACCG 1001 CCCCTACGTT GGCACTGCC CACAGTTGC CACCAGTTGC CACAGCCGC CTCAAAACCG 1001 CCCCTACGTT GGCACTGCC CACAGTTGC CACCAGTTGC CACAGCGCC CACAGTTGC CACACCGCC CACAGTTGC CACACCGCC CACAGTTGC CAC</seq>	a	087					
The following partial DNA sequence was identified in N. meningitidis <seq 281="" id="">:  a087.seq  1 ATGGGCGGTA AAACCTTTAT GCTGATGGCG GGCGGACGG GGCGATATT 51 TTTCCCCGGC CTGGCGGTGG CGGATTCATT GCGCGCGCGC GGCCATCATG 101 TAATTTGCT GGGCAGCAAG GATTCGATG GCGCGCGCC GGCCATCATG 102 TAATTTGCT TGCTCGAAAC GCTGGCGTT AAAGGCGCAT CGTGCCGCAA 103 TACGACATCC TGCTCGAAAC GCTGGCGTT AAAGGCGTGC GCGCAACGG 201 CATCAAACGC AAGCTGATGC TGCCGTTTAC TTTGTATCAA ACTGTCCGCG 203 AAGCGCAGCA GATTATCCGC AAACACCGTG TCGAGTGCGT CATCGGCTTC 301 GGCGGCTTCG TTACCTTTCC CGGCGGTTTG GCGCGCAAGT TATAGGCGT 301 GCCGATTGT ATTCACGAGC AAAACCCCGT GCGCGAAGT TATAGGCGT 301 ACCTGTCGCG CTGGCGAAG CGGGTTTGT ACGCTTTCC GAAAGCGTC 401 ACCTGTCGCG CTGGCCGAAC CGGCGTTTGT ACGCTTTTCC GAAAGCGTTC 451 AGCCACGAAG GCGGCTTGGT CGGCAACCCC GTCCGCGCCG ATATTAGCAA 501 CCTGCCCGTG CCTGCCGAAC GCTTCCAAGG GCGTGTAGGC CGTCTGAAAA 501 TTTTGGTGT CGGCGCAAC TTGGGCGGG ACGTTTTGAA CAAAACCGTA 601 CCGCAGGCAT TGGCTTTGCT GCCCGAACAT GCGCGCCC AGATGTACCA 601 CCAATCGGA CGGGCAAGC TGGCAACT GCGCGCCC AGATGTACCA 601 CCAATCGGA CGGGCAACC TGGCCGAATTA TACACCGCC 701 TGGGCGTGA AGCGGAATGC GTGGAATTTA TTACCACATAT GGTGTCCGCC 701 TGGGCGTGA AGCGGAATGC GTGGAATTTA TTACCACATAT GGTGTCCGCC 801 CGACTTGACG GCGGCGGAT TGGGTCCGT GCCGGCGCC TGACGATTGC 801 CGACTTGAC GCGGCGGAT TGGGTCCGT GTAGTGCCC TACCCTCCACG 801 CGGGATTGC TGTTCCCGCA AACCCACGCC GTTTTATGTC GCAGCGGAG 901 GCGGATTGC TGTTCCCGCA ACCCACGCC GTTTTATGTC GCAGCGGAG 901 GCGGGATTGC TGTTCCCGCA ACCCACGCC GTTTTATGTC GCAGGCGGA 901 GCGGGATTGC TGTTCCCGCA ACCCACGCC GTTTTATGT GCAGCGGAA 901 GCGGGATTGC TGTTCCCGCA ACCCACGCC GTTTTATGT GCAGCGGAA 901 GCGGGATTGC TGTTCCCGCA ACCCACGCC GTTTTATGT GCAGCGGAA 901 GCGGGATTGC TGTTCCCGCA ACCCACGCC GTTTTATGT GCAGCGGAA 901 GCGGGATTGC TGTTCCCGCA ACCCACGCC GTTTTATGG GCAGAGCAGC 1001 CCCCTACGTT GGCACTGCC CACAGCGC GTTTTATGG GCAGAGAAACC 1001 CCCCTACGTT GGCACTGCC CACAGTGCG ACGACAGCCC GCAAAAACC 1001 CCCCTACGTT GGCACTGCC CACAGTGCG ACGACAGCCC GCAAAAACC</seq>	3						
1 ATGGGCGTA AAACCTTTAT GCTGATGGCG GGCGAACGG GCGACATAT 51 TTTCCCCGC CTGGCGGTGG CGGATTCATT 51 TTTCCCCGCG CTGGCGGTGG CGGATTCATT 51 TAATTTGGCT GGGCAGCAAG GATTCATTG CGCGCCGGC GCCATCATG 101 TAATTTGGCT GGGCAGCAAG GATTCATTG AAGAGCGCAT CGTGCCGAA 151 TACGACATCC TGCTCGAAAC GCTGGCGATT AAAGGCGTGC GCGCAACGG 201 CATCAAACGC AAGCTGATGC TGCCGTTTAC TTTGTATCAA ACTGTCCGCG 251 AAGCGCAGCA GATTATCCGC AAACACCGTG TCGAGTGCGT CATCGGCTTC 301 GGCGGTTTG TACCTTTCC CGGCGGTTTG GCGCGCAAGT TATTAGGCGT 351 GCCGATTGTG ATTCACGAGC AAACACCGTG GCGCGAAGT TATTAGGCGT 401 ACCTGTCGCG CTGGGCGAAG CGGGTTTGT ACGCTTTTCC GAAAGCGTTC 451 AGCCACGAAG GCGGCTTGGT CGGCAACCCC GTCCGCCGC ATATTAGCAA 501 CCTGCCCGTG CCTGCCGAAC GCTTCCAAGG GCCTTAAACACGCC 401 CCGCAGCAT TGGCTTTGCT GCCGCAAAA CCGTTTTGAAAA 551 TTTTGGTGGT CGCGCGCAGT TTGGGCGCGG ACGTTTGAAAAACGTA 601 CCGCAGGCAT TGGCTTTGCT GCCGCAAAT GCGCGTCCGC ACAATCACAC 651 CCAATCGGA CGGGCAAGC TGGGCGAAT TACCGACACAC GCCCGACAAT TACCGACGCC 701 TGGGCGTGCA ACGGGAATGC GTGGAATTTA TTACCGATAT GGTGTCCGCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGGT TACCTCACG 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGC TACCGATTCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGCGCG TACCGATTCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGC TACCGATTCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGCG TACCGATTCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGCG TACCGATTCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGCG TACCGATTCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGGAA AACCCCGC GTTTATTGCCG TATCCTCACG 801 CGAGTTGACG CCATTAAACC GCCAACGCG GTTTTATGGC GTTATCCTCACG 801 CGCGTATGC CGATTGCC GCCAACGCG GTTTTATGGT GCAGCGGAG 901 GCGGATTGC CGATTGCC GCCAACGCG GTTTTATGGC GTTATCCTCACG 801 CAGCGAAAACG GCCGAAAAACG GCCAACGCG GTTTTATGGT GCAGCGGAG 901 GCGGATTGC CGATTGCC GCCAACGCG GTTTTATGGT GCAGCGGAG 901 GCGGATTGC CGATTGCC GCCAACGCG GTTTTATGGT GCAGCGGAAAACG 901 GCGGATTGC CGATTGCC GCCAACGCG GTTTTATGGT GCAGCGGAGAAACG 901 GCGGAAAAACG GCCGAACACCC GTTTTATGGT GCAGCGGAGAAACG 901 GCGGAAAACG GCCGAACACCC GACAACGCG GTTTTATGGT GCAGCGGAAAACG 901 GCGGAAAAACG GCGGAAAACG GCCAACGGC GACAACGCG GAAAAACG GCGAAAACG GCGAAAACG GCGA							
1 ATGGGCGTA AAACCTTTAT GCTGATGGCG GGCGAACGG GCGACATAT 51 TTTCCCCGC CTGGCGGTGG CGGATTCATT 51 TTTCCCCGCG CTGGCGGTGG CGGATTCATT 51 TAATTTGGCT GGGCAGCAAG GATTCATTG CGCGCCGGC GCCATCATG 101 TAATTTGGCT GGGCAGCAAG GATTCATTG AAGAGCGCAT CGTGCCGAA 151 TACGACATCC TGCTCGAAAC GCTGGCGATT AAAGGCGTGC GCGCAACGG 201 CATCAAACGC AAGCTGATGC TGCCGTTTAC TTTGTATCAA ACTGTCCGCG 251 AAGCGCAGCA GATTATCCGC AAACACCGTG TCGAGTGCGT CATCGGCTTC 301 GGCGGTTTG TACCTTTCC CGGCGGTTTG GCGCGCAAGT TATTAGGCGT 351 GCCGATTGTG ATTCACGAGC AAACACCGTG GCGCGAAGT TATTAGGCGT 401 ACCTGTCGCG CTGGGCGAAG CGGGTTTGT ACGCTTTTCC GAAAGCGTTC 451 AGCCACGAAG GCGGCTTGGT CGGCAACCCC GTCCGCCGC ATATTAGCAA 501 CCTGCCCGTG CCTGCCGAAC GCTTCCAAGG GCCTTAAACACGCC 401 CCGCAGCAT TGGCTTTGCT GCCGCAAAA CCGTTTTGAAAA 551 TTTTGGTGGT CGCGCGCAGT TTGGGCGCGG ACGTTTGAAAAACGTA 601 CCGCAGGCAT TGGCTTTGCT GCCGCAAAT GCGCGTCCGC ACAATCACAC 651 CCAATCGGA CGGGCAAGC TGGGCGAAT TACCGACACAC GCCCGACAAT TACCGACGCC 701 TGGGCGTGCA ACGGGAATGC GTGGAATTTA TTACCGATAT GGTGTCCGCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGGT TACCTCACG 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGC TACCGATTCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGCGCG TACCGATTCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGC TACCGATTCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGCG TACCGATTCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGCG TACCGATTCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGCG TACCGATTCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGGAA AACCCCGC GTTTATTGCCG TATCCTCACG 801 CGAGTTGACG CCATTAAACC GCCAACGCG GTTTTATGGC GTTATCCTCACG 801 CGCGTATGC CGATTGCC GCCAACGCG GTTTTATGGT GCAGCGGAG 901 GCGGATTGC CGATTGCC GCCAACGCG GTTTTATGGC GTTATCCTCACG 801 CAGCGAAAACG GCCGAAAAACG GCCAACGCG GTTTTATGGT GCAGCGGAG 901 GCGGATTGC CGATTGCC GCCAACGCG GTTTTATGGT GCAGCGGAG 901 GCGGATTGC CGATTGCC GCCAACGCG GTTTTATGGT GCAGCGGAAAACG 901 GCGGATTGC CGATTGCC GCCAACGCG GTTTTATGGT GCAGCGGAGAAACG 901 GCGGAAAAACG GCCGAACACCC GTTTTATGGT GCAGCGGAGAAACG 901 GCGGAAAACG GCCGAACACCC GACAACGCG GTTTTATGGT GCAGCGGAAAACG 901 GCGGAAAAACG GCGGAAAACG GCCAACGGC GACAACGCG GAAAAACG GCGAAAACG GCGAAAACG GCGA							
1 ATGGGCGTA AAACCTTTAT GCTGATGGCG GGCGAACGG GCGACATAT 51 TTTCCCCGC CTGGCGGTGG CGGATTCATT 51 TTTCCCCGCG CTGGCGGTGG CGGATTCATT 51 TAATTTGGCT GGGCAGCAAG GATTCATTG CGCGCCGGC GCCATCATG 101 TAATTTGGCT GGGCAGCAAG GATTCATTG AAGAGCGCAT CGTGCCGAA 151 TACGACATCC TGCTCGAAAC GCTGGCGATT AAAGGCGTGC GCGCAACGG 201 CATCAAACGC AAGCTGATGC TGCCGTTTAC TTTGTATCAA ACTGTCCGCG 251 AAGCGCAGCA GATTATCCGC AAACACCGTG TCGAGTGCGT CATCGGCTTC 301 GGCGGTTTG TACCTTTCC CGGCGGTTTG GCGCGCAAGT TATTAGGCGT 351 GCCGATTGTG ATTCACGAGC AAACACCGTG GCGCGAAGT TATTAGGCGT 401 ACCTGTCGCG CTGGGCGAAG CGGGTTTGT ACGCTTTTCC GAAAGCGTTC 451 AGCCACGAAG GCGGCTTGGT CGGCAACCCC GTCCGCCGC ATATTAGCAA 501 CCTGCCCGTG CCTGCCGAAC GCTTCCAAGG GCCTTAAACACGCC 401 CCGCAGCAT TGGCTTTGCT GCCGCAAAA CCGTTTTGAAAA 551 TTTTGGTGGT CGCGCGCAGT TTGGGCGCGG ACGTTTGAAAAACGTA 601 CCGCAGGCAT TGGCTTTGCT GCCGCAAAT GCGCGTCCGC ACAATCACAC 651 CCAATCGGA CGGGCAAGC TGGGCGAAT TACCGACACAC GCCCGACAAT TACCGACGCC 701 TGGGCGTGCA ACGGGAATGC GTGGAATTTA TTACCGATAT GGTGTCCGCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGGT TACCTCACG 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGC TACCGATTCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGCGCG TACCGATTCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGC TACCGATTCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGCG TACCGATTCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGCG TACCGATTCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGCG TACCGATTCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGGAA AACCCCGC GTTTATTGCCG TATCCTCACG 801 CGAGTTGACG CCATTAAACC GCCAACGCG GTTTTATGGC GTTATCCTCACG 801 CGCGTATGC CGATTGCC GCCAACGCG GTTTTATGGT GCAGCGGAG 901 GCGGATTGC CGATTGCC GCCAACGCG GTTTTATGGC GTTATCCTCACG 801 CAGCGAAAACG GCCGAAAAACG GCCAACGCG GTTTTATGGT GCAGCGGAG 901 GCGGATTGC CGATTGCC GCCAACGCG GTTTTATGGT GCAGCGGAG 901 GCGGATTGC CGATTGCC GCCAACGCG GTTTTATGGT GCAGCGGAAAACG 901 GCGGATTGC CGATTGCC GCCAACGCG GTTTTATGGT GCAGCGGAGAAACG 901 GCGGAAAAACG GCCGAACACCC GTTTTATGGT GCAGCGGAGAAACG 901 GCGGAAAACG GCCGAACACCC GACAACGCG GTTTTATGGT GCAGCGGAAAACG 901 GCGGAAAAACG GCGGAAAACG GCCAACGGC GACAACGCG GAAAAACG GCGAAAACG GCGAAAACG GCGA							
1 ATGGGCGTA AAACCTTTAT GCTGATGGCG GGCGAACGG GCGACATAT 51 TTTCCCCGC CTGGCGGTGG CGGATTCATT 51 TTTCCCCGCG CTGGCGGTGG CGGATTCATT 51 TAATTTGGCT GGGCAGCAAG GATTCATTG CGCGCCGGC GCCATCATG 101 TAATTTGGCT GGGCAGCAAG GATTCATTG AAGAGCGCAT CGTGCCGAA 151 TACGACATCC TGCTCGAAAC GCTGGCGATT AAAGGCGTGC GCGCAACGG 201 CATCAAACGC AAGCTGATGC TGCCGTTTAC TTTGTATCAA ACTGTCCGCG 251 AAGCGCAGCA GATTATCCGC AAACACCGTG TCGAGTGCGT CATCGGCTTC 301 GGCGGTTTG TACCTTTCC CGGCGGTTTG GCGCGCAAGT TATTAGGCGT 351 GCCGATTGTG ATTCACGAGC AAACACCGTG GCGCGAAGT TATTAGGCGT 401 ACCTGTCGCG CTGGGCGAAG CGGGTTTGT ACGCTTTTCC GAAAGCGTTC 451 AGCCACGAAG GCGGCTTGGT CGGCAACCCC GTCCGCCGC ATATTAGCAA 501 CCTGCCCGTG CCTGCCGAAC GCTTCCAAGG GCCTTAAACACGCC 401 CCGCAGCAT TGGCTTTGCT GCCGCAAAA CCGTTTTGAAAA 551 TTTTGGTGGT CGCGCGCAGT TTGGGCGCGG ACGTTTGAAAAACGTA 601 CCGCAGGCAT TGGCTTTGCT GCCGCAAAT GCGCGTCCGC ACAATCACAC 651 CCAATCGGA CGGGCAAGC TGGGCGAAT TACCGACACAC GCCCGACAAT TACCGACGCC 701 TGGGCGTGCA ACGGGAATGC GTGGAATTTA TTACCGATAT GGTGTCCGCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGGT TACCTCACG 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGC TACCGATTCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGCGCG TACCGATTCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGC TACCGATTCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGCG TACCGATTCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGCG TACCGATTCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGCG TACCGATTCC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GCCGGCGGAA AACCCCGC GTTTATTGCCG TATCCTCACG 801 CGAGTTGACG CCATTAAACC GCCAACGCG GTTTTATGGC GTTATCCTCACG 801 CGCGTATGC CGATTGCC GCCAACGCG GTTTTATGGT GCAGCGGAG 901 GCGGATTGC CGATTGCC GCCAACGCG GTTTTATGGC GTTATCCTCACG 801 CAGCGAAAACG GCCGAAAAACG GCCAACGCG GTTTTATGGT GCAGCGGAG 901 GCGGATTGC CGATTGCC GCCAACGCG GTTTTATGGT GCAGCGGAG 901 GCGGATTGC CGATTGCC GCCAACGCG GTTTTATGGT GCAGCGGAAAACG 901 GCGGATTGC CGATTGCC GCCAACGCG GTTTTATGGT GCAGCGGAGAAACG 901 GCGGAAAAACG GCCGAACACCC GTTTTATGGT GCAGCGGAGAAACG 901 GCGGAAAACG GCCGAACACCC GACAACGCG GTTTTATGGT GCAGCGGAAAACG 901 GCGGAAAAACG GCGGAAAACG GCCAACGGC GACAACGCG GAAAAACG GCGAAAACG GCGAAAACG GCGA	The fol	Howing porti	of DNIA comen	ce was identi	fied in M m	oninaitidia –	CEO ID 2015.
ATGGGCGGTA AAACCTTTAT GCTGATGGCG GGCGGAACGG GCGGACATAT TTTCCCCGCG CTGGCGGTGG CGGATTCATT GCGCGCGCGC GGCCATCATG GGCAGCAGCAG GATTCGATG AAGAGCCCAT CGTGCCGAA TAGACATCC TGCTCGAAAC GCTGGCGATT AAAGGCGTGC GCGCAACGG CATCAAACGC AAGCTGATGC TGCTGTTAC TTTGTATCAA ACTGTCCGCG AAGCGCAGCA GATTATCCGC AAACACCGTG TCGAGTGCGT CATCGGCTTC GGCGGCATCG TTACCTTTCC CGGCGGTTTG GCGGCGAACGG TATCACAGCC AAACACCGTG TCGAGTGCGT TATTAGGCGT GCGGATTGT ATTCACGAGC AAAACACCGTG TCCAACGCC CTGCCGAACG AAACACCGTG TCGAGTGCGT TATTAGGCGT ACCTGTCCGC CTGGCGAAC CGGCGTTTG ACCGTTTCT TCCAACCGCC CTGCCCGAAC GCGCCTTGGT CGCCAACCCC GTCCGCGCCG ATATTAGCAA GCCACCAAG GCGCTTGGT CGCCAACCCC GTCCGCGCC ATATTAGCAA CCTGCCCGTG CCTGCCGAAC GCTTCCAAGG GCGTGAAGGC CGTCTGAAAA TTTTTGGTGGT CGGCAGATC TTGGGCGCGG ACGTTTTGAA CAAAACCGTA GCCACGAGCAT TGGCTTTGCT GCCCGACAAT GCGCGTCCGC AGATGTACCA GCCACGAGCAT TGGCTTTGCT GCCCGACAAT GCGCGTCCGC AGATGACCGC TGGCGGGTA AGCGGAATGC TGGGAATTTA TTACCGAATAT GGTGTCCGCC TGGCGGTGC AGCGGATGC GTGGAATTTA TTACCGAATAT GGTGTCCGCC TACCGCGATG CCGATTTGGT GATTTGCCGT GCCGGCGGC TGACCGATTGC GCGGGTTGAC CCGATTTGGT GATTTGCCGT GCCGGCGGC TATCCTCACG GCGGGTTGAC CCGATTTGGT GATTTGCCGT GCCGGCGGC TATCCTCACG GCGGGTTGAC CCGATTTGGT GATTTGCCGT GCCGGCGGC TATCCTCACG GCGGGTTGAC CCGATTTGCT GATTTGCCG TGCCGGCGCG TATCCTCACG GCGGGTTGAC CCGATTTGGT GATTTGCCG TGCCGGCGCG TATCCTCACG GCGGGTTGAC CCGATTTGCT GCCGAACGCC GTTTTATTGGT GCCGGCGGAG GCGGGATTGC TGTTGCCGC AACCCAGTT ACGCCGCAAAACCC GCCAACGCG GTTTTATTGCGT GCCGGCAAAAACCC GCCAACGCG GTTTTATTGCGT GCCGGCAAAAACCC GCCAACGCG GCTTTAAACC GCCAACACCCC CTCAAATGG GCAGAAAACCC AAAACCCAGTT ACCCCACACCC CCCAACGCC CCCAACGCCC CCCAACGCCC CCCAACGCCC CCCAACGCCC CCCAACGCCC CCCAACGCCC CCCAACGCCC CCCAACGCCC CCCCAACGCCC CCCAACGCCC CCAACGCCC CCCAACGCCC CCCAACGCCC CCCAACGCCC CCCAACGCCC CCCAACGCCC CCCAACGCCC CCCAACGCCC CCCAACGCCC CCCAACGCCC CCCAACGCCC CCCAACGCCC CCCAACGCCC CCCAACGCCC CCCAACGCCC CCCAACGCCC CCCCAACGCCC CCCCAACGCCC CCCAACGCCCC CCCCAACGCCCC CCCAACGCCCC CCCCAACGCCC CCCC		~ .	iai DNA sequen	ce was ideiin	med m iv. me	eningiliais \	3EQ ID 2812:
TITCCCCGCG CTGGCGGTGG CGGATTCATT GCGCGCGCG GGCCATCATG  101 TAATTTGCT GGGCACAAG GATTCGATGC AAGAGCGCAT CGTGCCGCAA  151 TACGACATCC TGCTCGAAAC GCTGGCGATT AAAGGCGTGC GCGCAACAGG  201 CATCAAACGC AAGCTGATGC TGCCGTTTAC TTTGTATCAA ACTGTCCGCG  251 AAGCGCAGCA GATTATCCGC AAACACCGTG TCGAGTGCGT CATCGGCTTC  301 GGCGGCTTCG TTACCTTTCC CGGCGGTTTG GCGGCAACT TATTAGGCGT  351 GCCGATTGTG ATTCACGAGC AAACACCGTG TCCAACCGCC  401 ACCTGTCGCG CTGGGCGAAC CGGCTTTGT ACCTTTCC GAAAGCGTTC  451 AGCCACGAAG GCGCCTTGGT CGGCAACCCC GTCCGCCGC ATATTAGCAA  501 CCTGCCCGTG CCTGCCGAAC GCTTCCAAGG GCGTTTTCAACACCCC  451 AGCCACGAAG GCGCTTGGT GCCCAACCCC GTCCGCCGC ATATTAGCAA  551 TTTTGGTGGT CGGCGGCAAC GCTTCCAAGG GCGTTTGAAAAA  551 TTTTGGTGGT CGGCGCAAC TTGGCCGCG ACGTTTTGAA CAAAACCGTA  601 CCGCAGCAT TGGCTTTGCT GCCCGACACT GCAGCGCGC AGATGTACCA  651 CCAATCGGGA CGGGGCAAGC TGGGCAGCTT TACGACGCCC  751 TACCGCGATG CCGATTTGGT GATTTGCCGT GCCGGCGCG TGACGATTGC  801 CGAGTTGACG GCGCGGGAT TGGGTGCGT GTTAGTCCGC  851 CCGTTGATGA CCATCAAACC GCCAACACGC GTTTTATTGCT GCCGCGCG TTACCACCGC  851 CCGTTGATGA CCATCAAACC GCCAACACGC GTTTTATTGGT GCAGGCGGAG  901 GCGGGATTGC TGTTGCCGCA AACCCAGTTG GCCGCGAAAAACCG  1001 CCCGTACGTT GGCACTGCCC CACAGTGCG CCCCAAAAACCG  1001 CCCGTACGTT GGCACTGCCC CACAGTGCG CCCCAAAAACCG  1001 CCCGTACGTT GCCCGCAAAAACC CCCCAACACGCC CACAAAACCGCC CACAAAACCGCC CACAAAAACCG CCCCAAAAACCG CCCCAAAAAACCG CCCCAAAAAACCG CCCCAAAAAACCG CCCAAAAAACCG CCCCAAAAAACCG CCCAAAAAACCG CCCAAAAAACCG CCCAACACGCC CTTAAACCC CCCAACACCGC CTTAAACCC CCCAAAAAACCG CCCCAAAAAACCG CCCAAAAAACCG CCCCAAAAAACCG CCCCAAAAAACCG CCCAAAAAACCG CCCAAAA	a087.s	-					
101 TAATTTGGCT GGGCAGCAAG GATTCGATGG AAGAGCGCAT CGTGCCGCAA 151 TACGACATCC TGCTCGAAAC GCTGGCGATT AAAGGCGTGC GCGGCAACGG 201 CATCAAACGC AAGCTGATGC TGCCGTTTAC TTTGTATCAA ACTGTCCGCG 251 AAGCGCAGCA GATTATCCGC AAACACCGTG TCGAGTGCGT CATCGGCTTC 301 GGCGGCTTCG TTACCTTTCC CGGCGGTTTG GCGGCGAAGT TATTAGGCGT 351 GCCGATTGTG ATTCACGAGC AAAACGCCGT GCGGGCGAAGT TATTAGGCGT 401 ACCTGTCGCG CTGGCCGAAG CGGGTGTTGT ACGCTTTCC GAAAGCGCTC 451 AGCCACGAAG GCGGCTTGGT CGGCCAACCCC GTCCGCGCCG ATATTAGCAA 501 CCTGCCCGTG CCTGCCGAAC GCGCAACCCC GTCCGCGCCG ATATTAGCAA 551 TTTTGGTGGT CGCCGCACT TGGGCGCGG ACGTTTTGAA CAAAACCCTA 651 CCAATCGGAA TGGCTTTGCT GCCCGACAAT GCCGCTCCCC AGATGTACCA 651 CCAATCGGAA GCGGCAAGC TGGCAACTT TACCGCGCC 701 TGGGCGTGCA ACCGGAATGC GCCGCACATT TACCGCGCC 751 TACCGCGATG CCGGCAATCC GCCGCACATT TACCGCGCC 801 CGAGTTGACG GCGCAATGC GCGCAACTT TACCGCGCC TACCGCCC 801 CGAGTTGACG GCGCCAATGC GCGCGCGCC TACCGCCC 801 CGGGGATTG TGGTTCCGCC GCCGGCGCC TGACGATTGC 801 CGGGATTG CCGAACCCC GCCAACGCG GTTTAATGGC GCAGCGGAG 901 GCGGGATTG TGTTCCCCCA ACCCAGTTG GCCGGCGCAC TACCCCCGA 951 GATTCTCGCC GCCAACGCG GTTTTATGGT GCAGGCGGAG 901 GCGGGATTG TGTTCCCCCA ACCCAGTTG GCAGAAAACG 1001 CCCGTACGTT GGCACTCCC CACAGTGCG ACCACAGTG CCCCAACGCG GTTTTATGGT GCAGGCGGAG 1051 ATTGCGTGT CGCCGTAA							
TACGACATCC TGCTCGAAAC GCTGGCGATT AAAGGCGTGC GCGGCAACGG CATCAAACGC AAGCTGATGC TGCCGTTTAC TTTGTATCAA ACTGTCCGCG CATCAAACGC AAGCTGATGC TGCCGTTTAC TTTGTATCAA ACTGTCCGCG CATCAAACGC AAACACCGTG TCGAGTGCG CATCGGCTTC CATCAAACGC AAACACCGTG TCGAGTGCGT CATCGGCTTC CATCAACCGC CTGGCGCAAC CGGCGGTTTG GCGGCGAAGT TATTAGGCGC CATCAACCGC CTGGCGCAAC CGGTGTTGT ACCCTTTCC GAAAGCGTCC CATCACCAC CGGCGCAAC CGGCGCAACCC GTCCGCGCCG ATATTAGCAA CATCACCAC CTTCCCAAC CGTTCCAAACC GTCCGCGCCG ATATTAGCAA CATCACCAC CTTCCCAGAC CGTTCCAAGAG CGTCTGAAAA CATCACCAC CGCCGCAAC CGTTCCAAGAG CGTCTGAAAA CACCCTTTTACACCACCACAC CTTCCAAGAG CGTCTGAAAA CACCCTTTACCACACCACACAC CGCCGCACACT CGCCGCCGCACACACACCACAC							
201 CATCAAACGC AAGCTGATGC TGCCGTTTAC TTTGTATCAA ACTGTCCGCG 251 AAGCGCAGCA GATTATCCGC AAACACCGTG TCGAGTGCGT CATCGGCTTC 301 GGCGGCTTCG TTACCTTTCC CGGCGGTTTG GCGCGAAGT TATTAGGCGT 351 GCCGATTGTG ATTCACGAGC AAAACGCCGT GGCAGGTTTG TCCAACCGCC 401 ACCTGTCGCG CTGGCGAAG CGGGTGTTG ACGCTTTTCC GAAAGCGTTC 451 AGCCACGAAG GCGCTTGGT CGGCAACCCC GTCCGCCGC ATATTAGCAA 501 CCTGCCCGTG CCTGCCGAAC GCTTCCAAGG GCGTGAAGG CGTCTGAAAA 551 TTTTGGTGGT CGGCGAGT TTGGGCGCGG ACGTTTTGAA CAAAACCGTA 601 CCGCAGGCAT TGGCTTTGCT GCCCGACAAT GCGCGTCCGC AGATGTACCA 651 CCAATCGGGA CGGGGCAAGC TGGGCAGCT GCAGGCGGAT TACGACCGC 701 TGGGCGTGCA AGCCGAATGC GTGGAATTA TTACCGATAT GGTGTCCGCC 751 TACCGCGATG CCGATTTGGT GTGGAATTTA TTACCGATAT GGTGTCCGCC 851 CCGTTGATGA CCACCACCC GTCGCACCC TGACGATTGC 801 CGGGTTGAC CCGCCGCGAT TGGGTGCGT GTTAGTGCCG TATCCTCACCG 851 CCGTTGATGA CCACCACCC GTCGCACCC TACCGACCGC 901 GCGGGATTGC TGTTGCCGCA AACCCAGTTG GCAGCGGAAAACC 901 GCGGATTGC TGTTGCCGCA AACCCAGTTG ACGCCGGAAAAACC 1001 CCCGTACGTT GGCACTGCCG CACAGTGCG CCACAGCGC CCCAAAACCC 1051 ATTGCGTGTG CGGCGTAA							
AAGCGCAGCA GATTATCCGC AAACACCGTG TCGAGTGCGT CATCGGCTTC GGCGGCTTCG TTACCTTTCC CGGCGGTTTG GCGGCGAAGT TATTAGGCGT GCCGATTGTG ATTCACGAGC AAAACGCCGT GGCAGGTTTG TCCAACCGCC ACT ACCTGTCGCG CTGGGCGAAG CGGGTGTTGT ACGCTTTCC GAAAGCGTTC AS1 AGCCACGAAG GCGCTTGGT CGGCAACCCC GTCCGCGCCG ATATTAGCAA CCTGCCCGTG CCTGCCGAAC GCTTCCAAGG GCGTGAAGGC CGTCTGAAAA CCTGCCCGTG CCGCGCAC TTGGGCGCGG ACGTTTTGAA CAAAACCGTA CCGCAGGCAT TGGCTTTGCT GCCCGACAT GCGCTCCGC AGATGTACCA CCGCAGCAT TGGCTTTGCT GCCCGACAAT GCGCTCCGC AGATGTACCA CTGCCCGGAACC CGGGCAACT GCGCGCGCT TACGACCGCC CTGCCCGCACAAT TTACCCGAAT GCGCGCGCC TACGACCGCC CTGCCCGCC TTACGCCCC TTACCCCCC CCGAATCGCG CCGAATTTA TTACCGATAT GCTGTCCGCC CCGAATTGGT GCGCGCGTT TGGGTGCGT TGCGCGCCC TGACGATTGC CCGATTGACA CCGCGCGGAT TGGGTGCGTT GCAGGCGCGC TACCCCCCC CCGACGTTGACG CCGCCGGGAT TGGGTGCGTT GCAGGCGCGC TACCCTCACG CCGATTGACG CCGCCGGGAT TGGGTGCGTT GCAGGCGCGC TACCCCCCC CCGACGTTGATGA CCATCAAACC GCCAACGCCC GTTTATTGGT GCAGGCGGAG CCGTTGATGA CCATCAAACC GCCAACGCCC GTTTTATGGT GCAGGCGGAG CCGGGGATTGC TGTTGCCGCA AACCCAGTTG ACGCCGGAAAAACG CCCGTACGTT GCCCCCAACGCC GTTTTATGGT GCAGGCGGAG CCCCTAAATCG GCCAACACCCC CCCCCAACGCCC GTTTTATGGT GCAGGCGGAG CCCCTAAAACC GCCAAAAAACG CCCCCAACGCCC GTTTTATGGT GCAGGCGGAG CCCCTAAAACC GCCAAAAAACC CCCCAACGCCC GTTTTATGGT GCAGGCGGAG CCCCCTAAAACC GCCAAAAAACC GCCAAAAAACC GCCAAAAAACC GCCAAAAAACC GCCAAAAAACC CCCCCAACGCC GTTTTATGGT GCAGGCGCGAAAAACC GCCAAAAAACC CCCCCAACGCCC GTTTTATGGT GCAGGCGCGAAAAACC GCCAAAAAACC CCCCCAACGCCC GTTTTATGGT GCAGGCGCGAAAAACC GCCAAAAAACC CCCCCAACGCCC GTTTTATGGT GCAGGCGCGAAAAACC GCCAAAAAACC CCCCCAACGCCC GTTTTATGGT GCAGGCGCGAAAAACC GCCCAACGCCC GCAAAAAACC GCCAAAAAACC CCCCCAACGCCC GTTTTATGGT GCAGGCGCAAAAACC GCCCAACGCCC GCCAACGCCC GCCAACGCCC GCCAACGCCC GCAAAAAACC GCCAAAAAACC CCCCAACGCCC GCCAACGCCC GCCAACGCCC GCCAACGCCC GCAAAAAACC CCCCCAACCCCC GCCAACGCCC GCCAACCCCC GCAAAAAACC CCCCAACGCCC GCCAACGCCC GCCAACCCCC GCAAAAAACC CCCCAACGCCC GCCAACCCCC GCCAACCCCC GCAAAAAACC CCCCAACCCCC CACACGCCC CACACGCCC CCAACCCCC CCAACCCCCC CACACCCCC CACACGCCC CCACACCCCC CACACCCCC CACACCCCCC CCACACCCCC CCACACCCCC							
GCGGCTTCG TTACCTTTCC CGGCGGTTTG GCGGCGAAGT TATTAGGCGT  351 GCCGATTGTG ATTCACGAGC AAAACGCCGT GGCAGGTTTG TCCAACCGCC  401 ACCTGTCGCG CTGGGCGAAG CGGGTGTTGT ACGCTTTTCC GAAAGCGTTC  451 AGCCACGAAG GCGGCTTGGT CGGCAACCCC GTCCGCGCCG ATATTAGCAA  501 CCTGCCCGTG CCTGCCGAAC GCTTCCAAGG GCGTGAAGGC CGTCTGAAAA  551 TTTTGGTGGT CGGCGCAGT TTGGGCGCG ACGTTTTGAA CAAAACCGTA  601 CCGCAGGCAT TGGCTTTGCT GCCCGACAAT GCGCGTCCGC AGATGTACCA  651 CCAATCGGGA CGGGCAAGC TGGGCAGCTT GCAGCGCGT  701 TGGGCGTGCA AGCGGAATGC GTGGAATTTA TTACCGATAT GGTGTCCGCC  751 TACCGCGATG CCGATTTGGT GATTTGCCGT GCCGGCGCG TGACGATTGC  801 CGAGTTGACG GCGGCGGAT TGGGTGCGTT GTTAGTGCCG TATCCTCACG  851 CCGTTGATGA CCATCAAACC GCCAACGCC GTTTTATGGT GCAGGCGGAG  901 GCGGGATTGC TGTTGCCGCA AACCCAGTTG ACGGCGGAAA AACTCGCCGA  951 GATTCTCGCC GGCTTAAACC GCCAACGCC CTCTAAATGG GCAGAAAACG  1001 CCCGTACGTT GGCACTGCCG CACAGTCG CCTCAAATGG GCAGAAAACG  1001 CCCGTACGTT GGCACTGCCG CACAGTCG CCGAAGACCGC CGAAGACCGC  1051 ATTGCGTGTG CGGCGTAA							
351 GCCGATTGTG ATTCACGAGC AAAACGCCGT GGCAGGTTTG TCCAACCGCC 401 ACCTGTCGCG CTGGGCGAAG CGGGTGTTGT ACGCTTTTCC GAAAGCGTTC 451 AGCCACGAAG GCGGCTTGGT CGGCAACCCC GTCCGCGCCG ATATTAGCAA 501 CCTGCCCGTG CCTGCCGAAC GCTTCCAAGG GCGTGAAGGC CGTCTGAAAA 551 TTTTGGTGGT CGGCGCAGT TTGGGCGCGG ACGTTTTGAA CAAAACCGTA 601 CCGCAGGCAT TGGCTTTGCT GCCCGACAAT GCGCGTCCGC AGATGTACCA 651 CCAATCGGA CGGGGCAAGC TGGGCAGCTT GCAGGCGGAT TACGACGCGC 701 TGGGCGTGCA AGCGGAATGC GTGGAATTTA TTACCGATAT GGTTCCCCC 751 TACCGCGATG CCGATTTGGT GATTTGCCGT GCCGGCCGC TGACGATTGC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GTTAGTGCCG TATCCTCACG 851 CCGTTGATGA CCATCAAACC GCCAACGCGC GTTTTATGGT GCAGGCGGAG 901 GCGGGATTGC TGTTGCCGA AACCCAGTTG ACGCGGAAA AACTCGCCGA 951 GATTCTCGGC GGCTTAAACC GCGAAAAATG CCTCAAATGG GCAGAAAACG 1001 CCCGTACGTT GGCACTGCCG CACAGTCG CGAAGCCGCG 1051 ATTGCGTGTG CGGCGTAA							
401 ACCTGTCGCG CTGGGCGAAG CGGGTGTTGT ACGCTTTTCC GAAAGCGTTC 451 AGCCACGAAG GCGGCTTGGT CGGCAACCCC GTCCGGCCG ATATTAGCAA 501 CCTGCCCGTG CCTGCCGAAC GCTTCCAAGG GCGTGAAGGC CGTCTGAAAA 551 TTTTGGTGGT CGGCGGCAGT TTGGGCGCG ACGTTTTGAA CAAAACCGTA 601 CCGCAGGCAT TGGCTTTGCT GCCCGACAAT GCGCGTCCGC AGATGTACCA 651 CCAATCGGA CGGGGCAAGC TGGGCAGCTT GCAGGCGGAT TACGACGCC 701 TGGGCGTGCA AGCGGAATGC GTGGAATTTA TTACCGATAT GGTTCCGCC 751 TACCGCGATG CCGATTTGGT GATTTGCCGT GCCGGCGCC TGACGATTGC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GTTAGTGCCG TATCCTCACG 851 CCGTTGATGA CCATCAAACC GCCAACGCC GTTTTATGGT GCAGGCGGAG 901 GCGGGATTGC TGTTGCCGA AACCCAGTTG ACGCGGAAA AACTCGCCGA 951 GATTCTCGCC GGCTTAAACC GCGAAAAATG CCTCAAATGG GCAGAAAACG 1001 CCCGTACGTT GGCACTGCCG CACAGTCG CGAAGCCGCG 1051 ATTGCGTGTG CGGCGTAA						•	
AGCCACGAAG GCGGCTTGGT CGGCAACCCC GTCCGCGCCG ATATTAGCAA  501 CCTGCCCGTG CCTGCCGAAC GCTTCCAAGG GCGTGAAGGC CGTCTGAAAA  551 TTTTGGTGGT CGGCGGCAGT TTGGGCGCGG ACGTTTTGAA CAAAACCGTA  601 CCGCAGGCAT TGGCTTTGCT GCCCGACAAT GCGCGTCCGC AGATGTACCA  651 CCAATCGGGA CGGGGCAAGC TGGGCAGCTT GCAGGCGGAT TACGACGCGC  701 TGGGCGTGCA AGCGGAATGC GTGGAATTTA TTACCGATAT GGTGTCCGCC  751 TACCGCGATG CCGATTTGGT GATTTGCCGT GCCGGCGCGC TGACGATTGC  801 CGAGTTGACG CCGATCGAACC GCCAACGCCG GTTTTATGGT GCAGGCGGAG  901 GCGGGATTGC TGTTGCCGCA AACCCAGTTG ACGGCGGAAA AACTCGCCGA  951 GATTCTCGGC GGCTTAAACC GCCAACAGCG CTTTTATGGT GCAGGCGGAG  901 CCCGTACGTT GGCACTGCCG CACAGGCGG ACGACGTGC CGAAGACCG  1001 CCCGTACGTT GGCACTGCCG CACAGTGCG ACGACGTGC CGAAGCCGCG  1051 ATTGCGTGTG CGGCGTAAA							
501 CCTGCCGTG CCTGCCGAAC GCTTCCAAGG GCGTGAAGGC CGTCTGAAAA 551 TTTTGGTGGT CGGCGGCAGT TTGGGCGCGG ACGTTTTGAA CAAAACCGTA 601 CCGCAGGCAT TGGCTTTGCT GCCCGACAAT GCGCGTCCGC AGATGTACCA 651 CCAATCGGGA CGGGGCAAGC TGGGCAGCTT GCAGGCGGAT TACGACGCGC 701 TGGGCGTGCA AGCGGAATGC GTGGAATTA TTACCGATAT GGTGTCCGCC 751 TACCGCGATG CCGATTTGGT GATTTGCCGT GCCGGCGCGC TGACGATTGC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGT GTTAGTGCCG TATCCTCACG 851 CCGTTGATGA CCATCAAACC GCCAACGCGC GTTTTATGGT GCAGGCGGAG 901 GCGGGATTGC TGTTGCCGCA AACCCAGTTG ACGGCGGAAA AACTCGCCGA 951 GATTCTCGGC GGCTTAAACC GCGAAAAATG CCTCAAATGG GCAGAAAACG 1001 CCCGTACGTT GGCACTGCCG CACAGTGCG ACGACGTTGC CGAAGCCGCG 1051 ATTGCGTGTG CGGCGTAA							
551 TTTTGGTGGT CGGCGGCAGT TTGGGCGCGG ACGTTTTGAA CAAAACCGTA 601 CCGCAGGCAT TGGCTTTGCT GCCCGACAAT GCGCGTCCGC AGATGTACCA 651 CCAATCGGGA CGGGGCAAGC TGGGCAGCTT GCAGGCGGAT TACGACGCGC 701 TGGGCGTGCA AGCGGAATGC GTGGAATTTA TTACCGATAT GGTGTCCGCC 751 TACCGCGATG CCGATTTGGT GATTTGCCGT GCCGGCGCG TGACGATTGC 801 CGAGTTGACG GCGGCGGGAT TGGGTGCGT GTTAGTGCCG TATCCTCACG 851 CCGTTGATGA CCATCAAACC GCCAACGCGC GTTTTATGGT GCAGGCGGAG 901 GCGGGATTGC TGTTGCCGCA AACCCAGTTG ACGGCGGAAA AACTCGCCGA 951 GATTCTCGGC GGCTTAAACC GCGAAAAATG CCTCAAATGG GCAGAAAACG 1001 CCCGTACGTT GGCACTGCCG CACAGTGCG ACGACGTTGC CGAAGCCGCG 1051 ATTGCGTGTG CGGCGTAA							
601 CCGCAGGCAT TGGCTTTGCT GCCCGACAAT GCGCGTCCGC AGATGTACCA 651 CCAATCGGGA CGGGGCAAGC TGGGCAGCTT GCAGGCGGAT TACGACGCGC 701 TGGGCGTGCA AGCGGAATGC GTGGAATTTA TTACCGATAT GGTGTCCGCC 751 TACCGCGATG CCGATTTGGT GATTTGCCGT GCCGGCGCGC TGACGATTGC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGTT GTTAGTGCCG TATCCTCACG 851 CCGTTGATGA CCATCAAACC GCCAACGCGC GTTTTATGGT GCAGGCGGAG 901 GCGGGATTGC TGTTGCCGCA AACCCAGTTG ACGGCGGAAA AACTCGCCGA 951 GATTCTCGGC GGCTTAAACC GCGAAAAATG CCTCAAATGG GCAGAAAACG 1001 CCCGTACGTT GGCACTGCCG CACAGTGCG ACGACGTTGC CGAAGCCGCG 1051 ATTGCGTGTG CGGCGTAA							
651 CCAATCGGGA CGGGGCAAGC TGGGCAGCTT GCAGGCGGAT TACGACGCGC 701 TGGGCGTGCA AGCGGAATGC GTGGAATTTA TTACCGATAT GGTGTCCGCC 751 TACCGCGATG CCGATTTGGT GATTTGCCGT GCCGGCGCG TGACGATTGC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGTT GTTAGTGCCG TATCCTCACG 851 CCGTTGATGA CCATCAAACC GCCAACGCGC GTTTTATGGT GCAGGCGGAG 901 GCGGGATTGC TGTTGCCGCA AACCCAGTTG ACGGCGGAAA AACTCGCCGA 951 GATTCTCGGC GGCTTAAACC GCGAAAAATG CCTCAAATGG GCAGAAAACG 1001 CCCGTACGTT GGCACTGCCG CACAGTGCG ACGACGTTGC CGAAGCCGCG 1051 ATTGCGTGTG CGGCGTAA							
TGGGCGTGCA AGCGGAATGC GTGGAATTTA TTACCGATAT GGTGTCCGCC T51 TACCGCGATG CCGATTTGGT GATTTGCCGT GCCGGCGCGC TGACGATTGC B01 CGAGTTGACG GCGGCGGAT TGGGTGCGTT GTTAGTGCCG TATCCTCACG B51 CCGTTGATGA CCATCAAACC GCCAACGCGC GTTTTATGGT GCAGGCGGAG 901 GCGGGATTGC TGTTGCCGCA AACCCAGTTG ACGGCGGAAA AACTCGCCGA 951 GATTCTCGGC GGCTTAAACC GCGAAAAATG CCTCAAATGG GCAGAAAACG 1001 CCCGTACGTT GGCACTGCCG CACAGTGCGG ACGACGTTGC CGAAGCCGCG 1051 ATTGCGTGTG CGGCGTAA							
751 TACCGCGATG CCGATTTGGT GATTTGCCGT GCCGGCGCG TGACGATTGC 801 CGAGTTGACG GCGGCGGAT TGGGTGCGTT GTTAGTGCCG TATCCTCACG 851 CCGTTGATGA CCATCAAACC GCCAACGCGC GTTTTATGGT GCAGGCGGAG 901 GCGGGATTGC TGTTGCCGCA AACCCAGTTG ACGGCGGAAA AACTCGCCGA 951 GATTCTCGGC GGCTTAAACC GCGAAAAATG CCTCAAATGG GCAGAAAACG 1001 CCCGTACGTT GGCACTGCCG CACAGTGCGG ACGACGTTGC CGAAGCCGCG 1051 ATTGCGTGTG CGGCGTAA							
801 CGAGTTGACG GCGGCGGGAT TGGGTGCGTT GTTAGTGCCG TATCCTCACG 851 CCGTTGATGA CCATCAAACC GCCAACGCGC GTTTTATGGT GCAGGCGGAG 901 GCGGGATTGC TGTTGCCGCA AACCCAGTTG ACGGCGGAAA AACTCGCCGA 951 GATTCTCGGC GGCTTAAACC GCGAAAAATG CCTCAAATGG GCAGAAAACG 1001 CCCGTACGTT GGCACTGCCG CACAGTGCGG ACGACGTTGC CGAAGCCGCG 1051 ATTGCGTGTG CGGCGTAA							
851 CCGTTGATGA CCATCAAACC GCCAACGCGC GTTTTATGGT GCAGGCGGAG 901 GCGGGATTGC TGTTGCCGCA AACCCAGTTG ACGGCGGAAA AACTCGCCGA 951 GATTCTCGGC GGCTTAAACC GCGAAAAATG CCTCAAATGG GCAGAAAACG 1001 CCCGTACGTT GGCACTGCCG CACAGTGCGG ACGACGTTGC CGAAGCCGCG 1051 ATTGCGTGTG CGGCGTAA							
901 GCGGGATTGC TGTTGCCGCA AACCCAGTTG ACGCCGGAAA AACTCGCCGA 951 GATTCTCGGC GGCTTAAACC GCGAAAAATG CCTCAAATGG GCAGAAAACG 1001 CCCGTACGTT GGCACTGCCG CACAGTGCGG ACGACGTTGC CGAAGCCGCG 1051 ATTGCGTGTG CGGCGTAA							
951 GATTCTCGGC GGCTTAAACC GCGAAAAATG CCTCAAATGG GCAGAAAACG 1001 CCCGTACGTT GGCACTGCCG CACAGTGCGG ACGACGTTGC CGAAGCCGCG 1051 ATTGCGTGTG CGGCGTAA							
1001 CCCGTACGTT GGCACTGCCG CACAGTGCGG ACGACGTTGC CGAAGCCGCG 1051 ATTGCGTGTG CGGCGTAA							
1051 ATTGCGTGTG CGGCGTAA							
This components to the emine said sequence (SEO ID 202, ODE 007 ax.							
THIS CONCEDURES TO THE ARRINO ACID SCHEENCE SALLT HT 7877 (1877.8)	This co	orresponds to	the amino acid	sequence <	SEO ID 282:	ORF 087.a>	<b>&gt;:</b>
a087.pep					,	00,.00	•
1 MGGKTFMLMA GGTGGHIFPA LAVADSLRAR GHHVIWLGSK DSMEERIVPO	~~~.p		T.MA GGTGGHTFPA	TAVADSTRAP	CHHAIMICGR	DSMEERIUPO	
51 YDILLETLAI KGVRGNGIKR KLMLPFTLYQ TVREAQQIIR KHRVECVIGF							
101 GGFVTFPGGL AAKLLGVPIV IHEQNAVAGL SNRHLSRWAK RVLYAFPKAF							
151 SHEGGLVGNP VRADISNLPV PAERFQGREG RLKILVVGGS LGADVLNKTV	_						

	1111111 1111					
a087	MGGKTFMLMAGGTG	GHIFPALAVA	DSLRARGHHV	IWLGSKDSME	ERIVPQYDIL	LETLAI
	10	20	30	40	50	60

30

 ${\tt MGGKTFMLXXGGTGGHIFPALAVADSLRARGHHVIWLGSKDSMEERIVPQYGIRLETLAI}$ 

40

201 PQALALLPDN ARPQMYHQSG RGKLGSLQAD YDALGVQAEC VEFITDMVSA 251 YRDADLVICR AGA<u>LTIAELT AAGLGALLVP</u> YPHAVDDHQT ANARFMVQAE 301 AGLLLPQTQL TAEKLAEILG GLNREKCLKW AENARTLALP HSADDVAEAA

85.4% identity over a 355 aa overlap

20

351

m087/a087

m087.pep

IAÇAA*

	7.0						
-007	70	80	90	100	110	120	
m087.pep	KGVRGNGIKRKLML	PVTLYQTVR	EAQRIIRKHR\	ECVIGFGGFV	TFPGGLAAK	LLXVPIV	
-007		1 111111	]   ] :	111111111	111111111	11 1111	
a087	KGVRGNGIKRKLML	PFTLYQTVRI	EAQQIIRKHRV	ECVIGFGGFV	TFPGGLAAK	LLGVPIV	
	70 -	80	90	100	110	120	
	120						
m087.pep	130	140	150	160	170	180	
moo/.pep	IHEQNAVAGLSNRH	LSRWAKRVL	YAFPKAFSHEG	GLVGNPVRAD	ISNLPVPAER	RFQGREG	
a087	THEONAVACIONE!		<u> </u>	111111111	111111111	111111	
2007	IHEQNAVAGLSNRH 130	LSKWAKRVLY	AFPKAFSHEG		ISNLPVPAEF	RFQGREG	
	130	140	150	160	170	180	
	190	200				•	
m087.pep			210	220	230	240	
moo, pep	RLKILVVGGSLGAD	VLNKTVPHAL	ALLPDNARPH	MYHQSGRGKL	GILQAXXXXX	XXXXXX	
a087	PIKIL VVCCSI CAD		711111111:		1 111		
4007	RLKILVVGGSLGAD	ZOO	ALLPDNARPQ	MYHOSGRGKL		GVQAEC	
	190	200	210	220	230	240	
			250	0.50			
m087.pep	XX	vv		260	270	280	
The second second	••••		XXXXXXXAGL	JALLVPYPHAV	VDDHQTANAR	FMVQAE	
a087	VEFITDMVSAYRDAI	I.VTCRAGAI.	• • • • • • • • • • • • • • • • • • •	וווווווווווו		11111	
	250	260	270	280			
		200	270	200	290	300	
	290	300	310	320	330		
m087.pep	AGLLLPQTQLTAEKI		EKCLKWAENA	שבט אמגשם.דענדיי	33U 	70.37	
	1111111111111111	11111111			NAEAAIACA	AX.	
a087	AGLLLPQTQLTAEKI	AEILGGLNR	EKCLKWAENAF			1 l	
	310	320	330	340	350	HA.	
				0.0	330		
The following portial DNA							

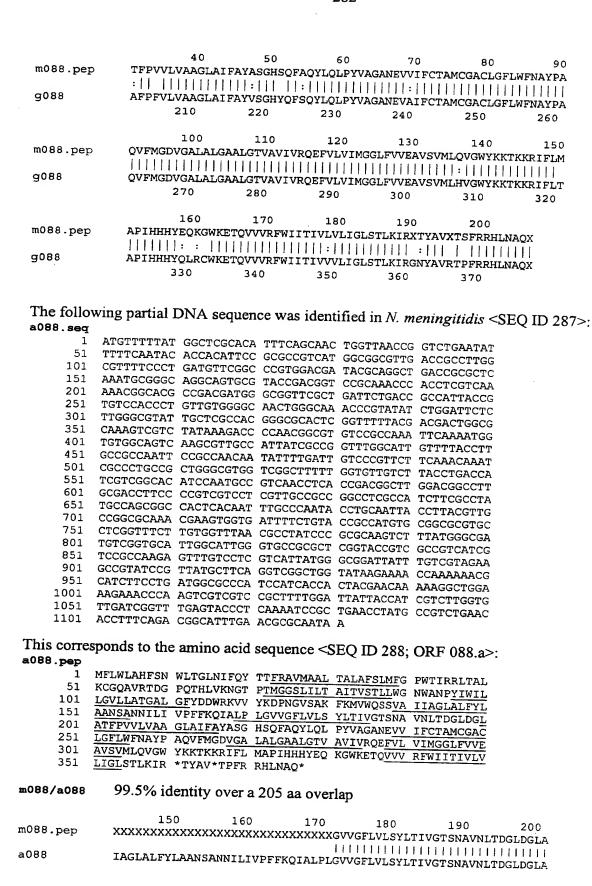
The following partial DNA sequence was identified in N.gonorrhoeae <SEQ ID 283>: g088.seq

```
ATGTTTTTAT GGCTCGCACA TTTCAGCAAC TGGTTAACCG GTCTGAATAT
   1
  51 TTTTCAATAC ACCACATTCC GCGCCGTTAT GGCGGCGTTG ACCGCCTTGG
 101 CGTTTTCCCT GATGTTCGGC CCGTGGACGA TACGCAGGCT GACCGCGCTC
 151 AAATGCGGGC AGGCAGTGCG TACCGACGGC CCGCAAACCC ACCTCGTCAA
 201 AAACGGCACG CCGACGATGG GCGGTTCGCT GATTCTGACC GCCATTACCG
 251 TGTCCACCCT GTTGTGGGGC AACTGGGCGA ACCCGTATAT CTGGATTCTC
 301 TTGGGCGTAC TGCTTGCCAC CGGTGCGCTC GGTTTTTACG ACGACTGGCG
 351 CAAAGTCGTT TATAAAGACC CCAACGGCGT GTCCGCCAAA TTCAAAATGG
401 TGTGGCAGTC AAGCGTTGCC GTTatcgcCG GTttggcaTT GTTTTACctt
 GCCCTGCCG CTGGGCGTGG TCGGCTTttt GGCCCGtttT TCAAACAAAT
TATTTTGATT GTCCCGtttT TCAAACAAAT
TCGCCTGCCG CTGGGCGTGG TCGGCTTttt GGtgttgTCT TACCTGACCA
 551 TCGTCGGCAC ATCCAACGCC GTCAACCTCA CcgaCGGCTT GGACGGCCTT
 601 GCCGCcttcc cgttcgtcct cgttgccgcC GGGCTCGCCA ttttcgccTA
 651 CGTCAGCGGA CACTACCAAT TTTCCCAATA CCTCCAGCTT CCCTATGTCG
 701 CCGGCGCGAA CGAAGTCGCT ATATTCTGCA CCGCCATGTG CGGCGCGTGC
 751 CTCGGATTTT TGTGGTTCAA CGCCTATCCC GCGCAAGTCT TTATGGGCGA
 801 TGTCGGCGCG CTGGCATTGG GTGCCGCGCT CGGTaccGtt gCCGTcaTcg
 851 tCCGCCAAGA ATTTGTcctc gtcattaTGG GCGGTCTGTT cgtcgtagaa
 901 gccgtgTCCG TTATGCTTCa tgtcggCTGG TACAAGAAAA Ccaaaaaaacg
      CATCTTcCTg acgGcaccga ttcatcacca ttaCCaactt cgatgCTGGa
     aagaaacgca agtcgtcgtc CGTTtCTGGA TTAtTAccat cgtcgtggtt
      tTgataggtt tGagtacccT caAAattcgc ggaaactatg ccgTCCGAAC
1101 ACCTTTCAGA CGGCATTTGA ACGCGCAATA A
```

This corresponds to the amino acid sequence <SEQ ID 284; ORF 088.ng>: g088.pep

- MFLWLAHFSN WLTGLNIFQY TTFRAVMAAL TALAFSLMFG PWTIRRLTAL 51
- KCGQAVRTDG PQTHLVKNGT PTMGGSLILT AITVSTLLWG NWANPYIWIL 101 LGVLLATGAL GFYDDWRKVV YKDPNGVSAK FKMVWQSSVA VIAGLALFYL
- 151 AANSANNILI VPFFKQIALP LGVVGFLVLS YLTIVGTSNA VNLTDGLDGL

```
201 AAFPFVLVAA GLAIFAYVSG HYQFSQYLQL PYVAGANEVA IFCTAMCGAC
        LGFLWFNAYP AQVFMGDVGA LALGAALGTV AVIVRQEFVL VIMGGLFVVE
        AVSVMLHVGW YKKTKKRIFL TAPIHHHYQL RCWKETQVVV RFWIITIVVV
        LIGLSTLKIR GNYAVRTPFR RHLNAQ*
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 285>:
m088.seq
       ATGTTTTAT GGCTCGCACA TTTCAGCAnC TGGTTAACCG GTCTGAATnn
       51
       101
    151
       201
       251
       301
       351
       501 nnnnnnnn nnnGGCGTGG TCGGCTTTTT GGTGTTGTCT TACCTGACCA
    551 TCGTCGGCAC ATCCAATGCC GTCAACCTCA CCGACGGCTT GGACGGCCTT
    601 GCGACCTTCC CCGTCGTCCT CGTTGCCGCC GGCCTCGCCA TCTTCGCCTA
    651 TGCCAGCGGC CACTCACAAT TTGCCCAATA CCTGCAATTA CCTTACGTTG
    701 CCGGCGCAAA CGAAGTGGTG ATTTTCTGTA CCGCCATGTG CGGCGCGTGC
    751 CTCGGTTTCT TGTGGTTTAA CGCCTATCCC GCGCAAGTCT TTATGGGCGA
    801 TGTCGGTGCA TTGGCATTGG GTGCCGCGCT CGGTACCGTC GCCGTTATCG
       TCCGCCAAGA GTTTGTCCTC GTCATTATGG GCGGATTATT TGTCGTAGAA
    901
       GCCGTATCCG TTATGCTTCA GGTTGGCTGG TATAAGAAAA CCAAAAAACG
       CATCTTCCTG ATGGCGCCCA TCCATCACCA CTACGAACAA AAAGGCTGGA
       AAGAAACCCA AGTCGTCGTC CGCTTTTGGA TTATTACCAT CGTCTTGGTG
       TTGATCGGTT TGAGTACCCT CAAAATCCGC TGAACCTATG CCGTCTGAAC
   1051
   1101 ATCTTTCAGA CGGCATTTGA ACGCGCAATA A
     1 MFLWLAHFSN WLTGLNIFOY TTFRAVMAAL TALAFSLMFG PWTIRRLTAL
    51 KCGOAVRTDG POTHLVKNGT PTMGGSLILT AITVSTLLWG NWANPYIWIL
   101 LGVLLATGAL GFYDDWRKVV YKDPNGVSAK FKMVWQSSVA VIAGLALFYL
   151 AANSANNILI VPFFKQIALP LGVVGFLVLS YLTIVGTSNA VNLTDGLDGL
   201 AAFPFVLVAA GLAIFAYVSG HYQFSQYLQL PYVAGANEVA IFCTAMCGAC
   251 LGFLWFNAYP AQVFMGDVGA LALGAALGTV AVIVRQEFVL VIMGGLFVVE
   301 AVSVMLHVGW YKKTKKRIFL TAPIHHHYQL RCWKETQVVV RFWIITIVVV
       LIGLSTLKIR GNYAVRTPFR RHLNAQ*
This corresponds to the amino acid sequence <SEQ ID 286; ORF 088>:
m088.pep
       MFLWLAHFSX WLTGLNXXXX XXXXXXXXX XXXXXXXXX XXXXXXXXX
    151 XXXXXXXXX XXXXXXXXX XGVVGFLVLS YLTIVGTSNA VNLTDGLDGL
       ATFPVVLVAA GLAIFAYASG HSQFAQYLQL PYVAGANEVV IFCTAMCGAC
       LGFLWFNAYP AQVFMGDVGA LALGAALGTV AVIVRQEFVL VIMGGLFVVE
       AVSVMLQVGW YKKTKKRIFL MAPIHHHYEQ KGWKETQVVV RFWIITIVLV
       LIGLSTLKIR XTYAVXTSFR RHLNAQ*
Computer analysis of this amino acid sequence gave the following results:
Homology with a predicted ORF from N.gonorrhoeae
ORF 088 shows 91.7% identity over a 205 aa overlap with a predicted ORF (ORF 088.ng)
from N. gonorrhoeae:
m088/q088
                                        10
                                               20
                                 GVVGFLVLSYLTIVGTSNAVNLTDGLDGLA
m088.pep
                                 IAGLALFYLAANSANNILIVPFFKQIALPLGVVGFLVLSYLTIVGTSNAVNLTDGLDGLA
a088
              150
                  160
                              170
                                      180
                                              190
                                                     200
```



	150	160	170	180	190	200
	210	220	230	240	250	260
m088.pep	TFPVVLVAAGLAIF.	AYASGHSQF/	AQYLQLPYVAG	SANEVVIFCT.	AMCGACLGFI	WFNAYPA
			F		1111111111	THEFT
a088	TFPVVLVAAGLAIF.	AYASGHSQF/	<b>AQYLQLPYVA</b> G	SANEVVIFCT.	AMCGACLGFL	WFNAYPA
	210	220	. 230	240	250	260
	270	280	290	300	310	320
m088.pep	QVFMGDVGALALGA	ALGTVAVIV	RQEFVLVIMGG	SLFVVEAVSV	MLQVGWYKKT	KKRIFLM
	111111111111			111111111	1111111111	1111111
a088	QVFMGDVGALALGA	ALGTVAVIV	RQEFVLVIMGG	SLFVVEAVSV	MLQVGWYKKT	KKRIFLM
	270	280	290	300	310	320
	330	340	350	360	370	
m088.pep	APIHHHYEQKGWKE	rqvvvrfwi:	TIVLVLIGLS	TLKIRXTYA	VXTSFRRHLN	XQX
		11111111		111111111	111 11111	111
a088	APIHHHYEQKGWKE	rqvvvrfwi:	TIVLVLIGLS	TLKIRXTYA	VXTPFRRHLN	XQA
	330	340	350	360	370	

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 289>: g089.seq

- 1 ATGCCGCCCA AAATCACGAA GAGCGGGTTT TGCAAACCGG CAATCGCGGC
- 51 GGCGGTCGCG CCGACATTCG TGCCTTTGCT GTCGTCGATG AATACCACGC
- CGTTTTTCTC GCCGATTTTT TCCACACGGT GCGGCAAGCC TTGGAAGGTT 101
- TTGACGTGTT CCAGCAATGC TTCGCGCGGC AAACCGACGG CCTCGCACAA
- 201 AGCCACGGCA GCCATAACGT TGGCGGCGTT GTGCAAACCT TGCAGCGGGA 251 TGTCTTGCGT AGAAATCAAA TCTTCATTGC CTTGTTTTAA ACAGCCCGTC
- 301 CCGCGTTCCA ACCAAAAATC GGCTTCGTGT TCCAAGGAAA ACCGTTTCAC
- 351 TTCACGCCCT GCCCGTTTCA TGGCGCGGCA GAACACGTCG TCCGCATTCA
- 401 AAACCTGCAC TCCATCGCCA CGGAAAATCT CGGCTTTGGT ATGCGCGTAG

This corresponds to the amino acid sequence <SEQ ID 290; ORF 089.ng>:

g089.pep

- 1 MPPKITKSGF CKPAIAAAVA PTFVPLLSSM NTTPFFSPIF STRCGKPWKV
- 51 LTCSSNASRG KPTASHKATA AITLAALCKP CSGMSCVEIK SSLPCFKOPV
- 101 PRSNQKSASC SKENRFTSRP ARFMARQNTS SAFKTCTPSP RKISALVCA*

The following partial DNA sequence was identified in N. meningitidis <SEO ID 291>: m089.seq

- ATGCCGCCCA AAATCACkAw GAGCGGATTT TGCAAACCGG CAATCGCGGC GGCAGTCGCG CCGACATTCG TGCCTTTGCT GTCGTCGATA AACACCACGC
- CGTTTTTCTC GCCGATTTTT TCCACGCGGT GCGGCAGGCC TTGGAAGGTT 101
- 151 TTGACGTGTT CGAGCAATGC TTCGCGCGAC AAACCGATGG CCTCACACAA
- 201 AGCCACKGCA GCCATGACGT TAGCGGCGTT GTGCAKACCT TGCAACGGWA
- 251 TGTCTTGCGT GACAATCAAA TCTTCATTGC CTTGTTTCAG GCGGCCTGTC
- 301 TCGCGTTCCA ACCAGAAATC AGCTTCGTGT TCCAACGAAA ACCATTTTAC
- 351 CTCGCGCCCG GCACGCTTCA TCGCGCGGCA GAACGCATCG TCCGCATTCA
- 401 AAACCTGCAC GCCGTCGCCA CGGAAAATCT TGGCTTTGGT ATGCGCATAG

This corresponds to the amino acid sequence <SEQ ID 292; ORF 089>:

m089.pep

- MPPKITXSGF CKPAIAAAVA PTFVPLLSSI NTTPFFSPIF STRCGRPWKV 1
- LTCSSNASRD KPMASHKATA AMTLAALCXP CNGMSCVTIK SSLPCFRRPV
- SRSNQKSASC SNENHFTSRP ARFIARQNAS SAFKTCTPSP RKILALVCA*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae

ORF 089 shows 88.6% identity over a 149 aa overlap with a predicted ORF (ORF 089.ng) from N. gonorrhoeae:

m089/g089

m089.pep	10 MPPKITXSGFCKPA              MPPKITKSGFCKPA 10	1111111		! ! ! ! ! ! ! ! ! ! ! !	1:1111111	11111
m089.pep	70 KPMASHKATAAMTLI          :   KPTASHKATAAITLI 70			: :	1111111111	. 1 1 1 1 1
m089.pep g089	130 ARFIARQNASSAFKT    :    :      ARFMARQNTSSAFKT	НИН	111111			

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 293>: a089.seq

1	ATGCCGCCTA	AAATCACGAA	GAGCGGATTT	TGCAAACCGG	CAATCGCGGC
51	GGCGGTCGCA	CCGACGTTCG	TGCCTTTGCT	GTCGTCGATC	AACACCACCC
101	CATTTTTTTTCTC	CCCCAMMMM	FGG3.00000	or our comit	AACACCACGC
	CATITICIC	GCCGATTTTT	TCCACGCGGT	GCGGCAGGCC	TTGAAAGGTT
151	TTGACGTGTT	CGAGCAATGC	TTCGCGCGGC	AAACCGACGG	CTTCGCACAA
201	GGCAACGGCA	GCCATCACGT	TACTCCCCOTT	CECCTTCCC	OII COCHCAA
	CCCILICOCA	GCCATCACGI	INGIGGCGIT	GIGCAAGCCT	TGCAGCGGAA
251	TATCTTGCGT	GGCAATCAAA	TCTTCATTGC	CTTGTTTCAG	GCGACCTGTC
301	TCACGTTCCA	ACCAAAAATC	GCCTTCCTAT	TOODACCAAA	7.001.001.01.0
251	00000000		COCTICGIAL	ICCAACGAAA	ACCATTCAC
351	CTCGCGCCCG	GCGCGCTTCA	TCGCACGACA	GAACGCATCG	TCCGCATTCA
401	AAACCTGCAC	ACCGTCGCCA	CCCAAAAmom	MCC CMMM C CM	
		cccrcdcca	COGNAMAICI	1 GGCTTTGGT	ATGCGCGTAG

### This corresponds to the amino acid sequence <SEQ ID 294; ORF 089.a>: a089.pep

1 MPPKITKSGF CKPAIAAAVA PTFVPLLSSM NTTPFFSPIF STRCGRP*KV 51 LTCSSNASRG KPTASHKATA AITLVALCKP CSGISCVAIK SSLPCFRRPV 101 SRSNQKSASY SNENHFTSRP ARFIARQNAS SAFKTCTPSP RKILALVCA*

#### 91.9% identity over a 149 aa overlap m089/a089

	10	20	30	40	50	60
m089.pep	MPPKITXSGFCKPA	IAAAVAPTF	/PLLSSINTT	PEESPIESTRO	CD DWRIIT MAG	00
	111111 111111	111111111		1111111111	GVENVATICE	SNASRD
<b>a08</b> 9	MPPKITKSGFCKPA			. , , , , , , , , , , , , , , , , , , ,		
	10	20	30	40		
	<b></b> 0	20	30	40	50	60
	70	80	90	100	110	
m089.pep	KPMASHKATAAMTI.			100	110	120
1 -1-	KPMASHKATAAMTL	· I I I I I I I	PCATTV22F	CERRPVSRSN	OKSASCSNEN	NHFTSRP
a089			111:1111		11111 1111	11111
4007	KPTASHKATAAITLVALCKPCSGISCVAIKSSLPCFRRPVSRSNQKSASYSNENHFTSRP					
	70	80	90	100	110	120
	130	140	150			
m089.pep	ARFIARONASSAFK					
r-r	ARFIARQNASSAFKTCTPSPRKILALVCAX					
a089	ARFIARQNASSAFKTCTPSPRKILALVCAX					
4005						
	130	140	150			

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 295>: g090.seq

ATGCGCGTAG TCGAGCAAAT CGTCGTAGCG GTCGAGATGG TCTTCGGAAA

.. 👊

```
51 TGTTCATCAC CGTCGCCGCA GTCGGGCGCA GGCTTTCGGT GTTTTCCAGT
         101
              TGGAAGCTGG AAAGCTCcca CACCCACACG TCCGCCTTTT TGCCTTCgcg
             ctgCAATtct gcctccaaga cgggcgtacc gatATTGCCC GCAATGAcgg
         201 tatccagccc gcacttgatg CAGAGatagc ggaccaggct ggttaccgTG
         251 GTTttgccgt tgctgCcggt aatcgCaatc accttgtcgC CGCGGCGGtt
         301 cAcaaTGTCc gccaGCAATt ggATGTCGCC TAgCACGCGC .ccgccgTTT
         351 TGCttga
This corresponds to the amino acid sequence <SEQ ID 296; ORF 090.ng>:
    q090.pep
              MRVVEQIVVA VEMVFGNVHH RRRSRAQAFG VFQLEAGKLP HPHVRLFAFA
           1
              LQFCLQDGRT DIARNDGIQP ALDAEIADQA GYRGFAVAAG NRNHLVAAAV
              HNVRQQLDVA XHAXRRFA*
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 297>:
    m090.seq
              ATGCGCATAG TCGAGCAAGT CGTCGTAGCG GTCGAGATGG TCTTCGGAAA
              TGTTCAGCAC CGTCGCCGCA GTCGGACGCA GGCTTTCGGT GTTTTCCAGT
           51
         101
              TGGAAGCTGG AAAGCTCCAA CACCCACACG TCCGCCTTTT TGCCTTCGCG
              CTGCCATTCC GCCTCCAAAA CCGGCGTGCC GATATTGCCC GCGATAACGG
              TATCCAGCCC GCACTTGATA CAGAGATAGC CGACCAGGCT CGTTACCGTG
         251 GTTTTGCCGT TGCTGCCGGT AATCGCAATT ACCTTGTCGT CCCGGCGGTT
         301 CACAATGTCC GCCAGCAATT CGATGTCGCC CAACACGCGT .CCGCCGTTT
              TGCTTGA
         351
This corresponds to the amino acid sequence <SEQ ID 298; ORF 090>:
    m090.pep
              MRIVEOVVVA VEMVFGNVOH RRRSRTOAFG VFOLEAGKLO HPHVRLFAFA
              LPFRLQNRRA DIARDNGIQP ALDTEIADQA RYRGFAVAAG NRNYLVVPAV
         101 HNVRQQFDVA QHAXRRFA*
Computer analysis of this amino acid sequence gave the following results:
Homology with a predicted ORF from N. gonorrhoeae
ORF 090 shows 83.9% identity over a 118 aa overlap with a predicted ORF (ORF 090.ng)
from N. gonorrhoeae:
    m090/g090
                                  20
                                            30
                                                      40
                 MRIVEQVVVAVEMVFGNVQHRRRSRTQAFGVFQLEAGKLQHPHVRLFAFALPFRLQNRRA 🕖
    m090.pep
                 MRVVEQIVVAVEMVFGNVHHRRRSRAQAFGVFQLEAGKLPHPHVRLFAFALQFCLQDGRT
    g090
                                  20
                                            30
                                                      40
                         10
                         70
                                   80
                                            90
                                                     100
                                                              110
                 DIARDNGIQPALDTEIADQARYRGFAVAAGNRNYLVVPAVHNVRQQFDVAQHAXRRFAX
     m090.pep
                 DIARNDGIQPALDAEIADQAGYRGFAVAAGNRNHLVAAAVHNVRQQLDVAXHAXRRFAX
     g090
                         70
                                  80
                                            90
                                                     100
                                                              110
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 299>:
a090.seq
         ATGCGCGTAG TCGAGCAAGT CGTCGTAGCG GTCGAGATGG TCTTCGGAAA
      1
     51 TGTTCAGCAC TGTCGCCGCA GTCGGGCGCA GGCTTTCGGT GTTTTCCAGT
    101 TGGAAACTGG AAAGCTCCAA CACCCACACG TCCGCCTTTT TGCCTTCGCG
```

This corresponds to the amino acid sequence <SEQ ID 300; ORF 090.a>:

CTGCAATTCC GCCTCCAAAA CCGGCGCGCC GATATTGCCC GCGATAACGG

TATCCAGCCC ACACTTGATG CAGAGATAGC CGACCAGGCT CGTTACCGTG 251 GTTTTGCCGT TGCTGCCGGT AATCGCAATC ACCTTGTCGC CGCGGCGGTT 301 CACAATGTCC GCCAGCAATT CGATGTCGCC CAACACGCGT C.CGCCGTTT

201

351 CGCTTAA

```
a090.pep
        MRVVEQVVVA VEMVFGNVQH CRRSRAQAFG VFQLETGKLQ HPHVRLFAFA
        LQFRLQNRRA DIARDNGIQP TLDAEIADQA RYRGFAVAAG NRNHLVAAAV
     51
    101
        HNVRQQFDVA QHAXRRFA*
m09/a090
           91.5% identity over a 117 aa overlap
                          20
                                   30
                                           40
                                                    50
           {\tt MRIVEQVVVAVEMVFGNVQHRRRSRTQAFGVFQLEAGKLQHPHVRLFAFALPFRLQNRRA}
m090.pep
           MRVVEQVVVAVEMVFGNVQHCRRSRAQAFGVFQLETGKLQHPHVRLFAFALQFRLQNRRA
a090
                  10
                          20
                                           40
                                                    50
                  70
                          80
                                   90
                                          100
                                                   110
                                                          119
           DIARDNGIQPALDTEIADQARYRGFAVAAGNRNYLVVPAVHNVRQQFDVAQHAXRRFAX
m090.pep
           a090
           DIARDNGIQPTLDAEIADQARYRGFAVAAGNRNHLVAAAVHNVRQQFDVAQHAXRRFAX
                          80
                                          100
```

The following partial DNA sequence was identified in N. gonorrhoeae g090-1.seq This sequence contains multiple stop codons (not shown)

This corresponds to the amino acid sequence < ORF 090-1.ng>: g090-1.pep (not shown)

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 2>: m090-1.seq

```
1 ATGACGGCGT TTGCATTTCA GACGGCATCA CAAAGCCTTA AACGCTTCGA
      TAAACACTTC CGAACGGTGC GCGTAGCCTT TGAACATATC AAAGCTCGCG
      CAGGCGGGGC TGAGCAACAC AATATCGCCT GCTTCGGCTT GGGCATATGC
 101
      CGTCTGAACG GCTTCTCCCA AAGTGGCGCA GTCGGTCATA TTCAAGCCGC
 151
      AGCCGTCCAA ATCGCGGCGG ATTTGCGGCG CATCGACACC AATCAAGAAC
 201
 251
      ACGCCTTTTG CCTTGCCTAC CAGTGCATCG CGCAGGGGCG TGAAGTCCTG
      CCCTTTACCC ATGCCGCCCA AAATCACGAA GAGCGGATTT TGCAAACCGG
 301
 351
      CAATCGCGGC GGCAGTCGCG CCGACATTCG TGCCTTTGCT GTCGTCGATA
      AACACCACGC CGTTTTTCTC GCCGATTTTT TCCACGCGGT GCGGCAGGCC
 401
 451
      TTGGAAGGTT TTGACGTGTT CGAGCAATGC TTCGCGCGAC AAACCGATGG
 501
      CCTCACACAA AGCCACGGCA GCCATGACGT TAGCGGCGTT GTGCAGACCT
 551
      TGCAACGGAA TGTCTTGCGT GACAATCAAA TCTTCATTGC CTTGTTTCAG
 601
      GCGGCCTGTC TCGCGTTCCA ACCAGAAATC AGCTTCGTGT TCCAACGAAA
 651
      ACCATTTAC CTCGCGCCCG GCACGCTTCA TCGCGCGGCA GAACGCATCG
      TCCGCATTCA AAACCTGCAC GCCGTCGCCA CGGAAAATCT TGGCTTTGGT
 701
 751
      ATGCGCATAG TCGAGCAAGT CGTCGTAGCG GTCGAGATGG TCTTCGGAAA
      TGTTCAGCAC CGTCGCCGCA GTCGGACGCA GGCTTTCGGT GTTTTCCAGT
 801
 851
      TGGAAGCTGG AAAGCTCCAA CACCCACACG TCCGCCTTTT TGCCTTCGCG
 901
      CTGCCATTCC GCCTCCAAAA CCGGCGTGCC GATATTGCCC GCGATAACGG
      TATCCAGCCC GCACTTGATA CAGAGATAGC CGACCAGGCT CGTTACCGTG
 951
1001
      GTTTTGCCGT TGCTGCCGGT AATCGCAATT ACCTTGTCGT CCCGGCGGTT
      CACAATGTCC GCCAGCAATT CGATGTCGCC CAACACGCGT CCGCCGTTTT
1051
      GCTTGAACGC CTCAATATCC GGCTGCCGCT CGCTGATGCC GGGACTGAGA
1101
     GCCAGAATAT CGAAACCGTT GTCCAGCGCA TCTTTCAGAC GGCCCGTGTA
1151
     AAACACCAAC CCGTCAAACA TCTTACCGAT TTGCGACACG CGTTCCGGCT
1201
1251
      TCAGCTCCGC ATCATACGCA GCAACCTCCG CGCCGTTTTT GCGCAGGTAG
     GCAATCATGG AAATACCCGT ACCGCCGAGT CCGGCGACGA GGATTTTTTT
1351
     GTTTTGAAAA GTCATTTTGG TTTGTCCTAA
```

This corresponds to the amino acid sequence <SEQ ID 3; ORF 090-1>: m090-1.pep

1	MTAFAFQTAS	QSLKRFDKHF	RTVRVAFEHI	KARAGGAEOH	NIACEGIGIC
51	RLNGFSQSGA	VGHIQAAAVQ	IAADLRRIDT	NOTHATCLAY	OCTAOCREVI
101	PFTHAAQNHE	ERILQTGNRG	GSRADIRAFA	VVDKHHAVFI.	ADEFUAURON
151	LEGFDVFEQC	FARQTDGLTQ	SHGSHDVSGV	VOTLORNVI.R	DNOTETALEO
201	AACLAFQPEI	SFVFQRKPFY	LAPGTLHRAA	ERIVATIONAL	AVATENTOEC
251	MRIVEQVVVA	VEMVFGNVQH	RRRSRTOAFG	VFOLEAGKT.O	HPHURT EN EN
301	LPFRLQNRRA	DIARDNGIQP	ALDTEIADOA	RYRGFAVAAG	NEMVI WORK
351	HNVRQQFDVA	QHASAVLLER	LNIRLPLADA	GTESONTETY	VODTEOTADI
401	KHQPVKHLTD	LRHAFRLQLR	IIRSNLRAVF	AQVGNHGNTR	TAESGDEDEE
451	VLKSHFGLS*				

```
The following partial DNA sequence was identified in N. gonorrhoeae <SEO ID 303>:
q091.seq
         ATGGAAATAC CCGTGCCGCC AAGTCCGGCG ACGAGGATTT TTTTGTTTGA
      51 AAGTCATTTT GGTTTTGTCC TAAAACAAAT CATATTGGGC AGGAGACGTC
     101 CGCCCTTGCC CAAGCCGCTT TCAGACGGCA TCGCGAGCCG ATTAATAACC
     151 CGCCTTCAGG CGTTGGTCAT TGTCGCAGCT GTTTTGGTCT CCGTTTTGAC
     201 AAGCCTTGCC AAGCCATTGT TGAGCGAGCG CAAGGTCTTG GCGCACGCCG
     251 CGTCCATCGT AATACATCAA GCCCAAATTG TATTGGGCTT GGGCATCCCC
     301 TTGTTCTGA
This corresponds to the amino acid sequence <SEQ ID 304; ORF 091.ng>:
g091.pep
      1
         MEIPVPPSPA TRIFLFESHF GFVLKQIILG RRRPPLPKPL SDGIASRLIT
      51 RLQALVIVAA VLVSVLTSLA KPLLSERKVL AHAASIVIHQ AQIVLGLGIP
     101 LF*
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 305>:
m091.seq
         ATGGAAATAC CCGTACCGCC GAGTCCGGCG ACGAGGATTT TTTTGTTTGA
      1
      51 AAAGTCATTT TGGTTTGTCC TAAAACAAAT CATATTGAGC AGGAGATGTC
     101 CGCCCTGCC CAAGCCGCTT TCAGACGGCA TCGCGAGCTG TTCAATAACC
     151 CGCCTTCAGG CGTTGGTCAT TGTCGCAGCC GTCTTGGTCT CCGTTTTGAC
     201 AAGCCTTGCC AAACCATTCT TGTGCAAGGG CGCGGTCTTG GCGCACGCCG
     251 CGTCTTTCGG CATACATCAC GCCCAAATTG TTTTGGGCTT GGGCTACCCC
     301 CTGCGC...
This corresponds to the amino acid sequence <SEQ ID 306; ORF 091>:
m091.pep
         MEIPVPPSPA TRIFLFEKSF WFVLKQIILS RRCPPLPKPL SDGIASCSIT
      51 RLQALVIVAA VLVSVLTSLA KPFLCKGAVL AHAASFGIHH AQIVLGLGYP
     101 LR.
Computer analysis of this amino acid sequence gave the following results:
Homology with a predicted ORF from N. gonorrhoeae
ORF 091 shows 84.2% identity over a 101 aa overlap with a predicted ORF (ORF 091.ng)
from N. gonorrhoeae:
m091/q091
                    10
                             20
                                       30
                                                40
                                                          50
                                                                   60
            MEIPVPPSPATRIFLFEKSFWFVLKQIILSRRCPPLPKPLSDGIASCSITRLQALVIVAA
m091.pep
            g091
            MEIPVPPSPATRIFLFESHFGFVLKQIILGRRRPPLPKPLSDGIASRLITRLQALVIVAA
                    10
                             20
                                       30
                                                40
                                                         50
                                                                   60
                    70
                             80
m091.pep
            VLVSVLTSLAKPFLCKGAVLAHAASFGIHHAQIVLGLGYPLR
            VLVSVLTSLAKPLLSERKVLAHAASIVIHQAQIVLGLGIPLFX
q091
                    70
                             80
                                       90
                                               100
The following partial DNA sequence was identified in N. meningiditis <SEO ID 307>:
a091.seq
         GAAATCATTT TGGTTTGTCC TAAAACAAAT CATATTGAGC AGGGGATGTC
     51
    101 TGATCCTGCT CAAGCCGCTT TCAGACGGCA TCGCGAGCTG TTCAATAACC
```

151 CGCTTTCAGG CGTTGGTCAT TGTCGCAGCT GTCTTGGTAT CCGTTTTGAC
201 AAGCCTTGCC AAGCCATTCT TGTGCAAGGG CGCGGTCTTG GCGCACGCCG

CGTCTTTCGG CATACATCAC GCCCAAATTG TTTTGGGC

BNSDOCID: <WO___9957280A2_i_>

251

This corresponds to the amino acid sequence <SEQ ID 308; ORF 091.a>: a091.pep

1 MEIPVPPSPA TRIFLFWKSF WFVLKQIILS RGCLILLKPL SDGIASCSIT 51 RFQALVIVAA VLVSVLTSLA KPFLCKGAVL AHAASFGIHH AQIVLG

```
m091/a091 93.8% identity over a 96 aa overlap
```

```
10
                          20
                                   30
                                           40
                                                   50
                                                            60
           MEIPVPPSPATRIFLFEKSFWFVLKQIILSRRCPPLPKPLSDGIASCSITRLQALVIVAA
m091.pep
           a091
           MEIPVPPSPATRIFLFWKSFWFVLKQIILSRGCLILLKPLSDGIASCSITRFQALVIVAA
                 10
                          20
                                  30
                                           40
                                                   50
                 70
                          80
                                  90
                                          100
           VLVSVLTSLAKPFLCKGAVLAHAASFGIHHAQIVLGLGYPLR
m091.pep
           ***********************
a091
           VLVSVLTSLAKPFLCKGAVLAHAASFGIHHAQIVLG
                          80
                                  90
```

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 309>:

```
ATGTTTTTTA TTTCAATCCG CTATATATTT GTCAGAAAAC TATGGTGCGC
   1
      AAACGGTCAG ACCTTTAAAA TAACGCCTTT ACGCACTAAA AACCAACCGG
   51
      AACGCAACAT TATGATGAAA AATCGAGTAA GCAACATCCA TTTTGTCGGT
  101
      ATCGGCGGCG TCGGCATGAG CGGTATCGCC GAAGTCTTGC ACAATTTGGG
      CTTTAAAGTT TCCGGTTCGG ATCAGGCGCG AAATGCCGCT ACCGAGCATT
      TGAGCAGCCT GGGCATTCAA GTTTATCCCG GCCATACCGC AGAACACGTT
  251
 301 AACGGTgcgg ATGTCGTCGT TGCCTCTACC GCCGTCAAGA AAGAAaatcC
 351 CGAAGTEGEC GCTGCGTTGG AGCGGCAAAT TCCCGTTATT CCGCGCGCCCT
 401 TGATGCTGGC AGAGCTGATG CGCTTCCGTG ACGGCatcgc cattgccggT
 451 ACGCACGGCA AAACCACGAC CACCAGCCTG ACCGCCTCCA TCCTCGGCGC
 501 GGCAGGACTC GACCCCACTT TCGTTATCGG CGGCAAACTC AACGCCGCAG
 551 GCACCAACGC CCGCTTGGGC AAAGGCCGAAT ACATCGTTGC CGAAGCCGAC
 601 GAATCCGATG CCTCTTTCCT ACATCTGACC CCGATTATGT CCGTCGTTAC
      CAATATCGAC GAAGACCATA TGGATACCTA CGGGCACAGC GTCGAAAAAC
 651
      TGCATCAGGC GTTTATCGAT TTCATCCACC GTATGCCCTT CTACGGCAAA
 701
 751 GCCTTTTTGT GTGTTGACAG CGAACACGTC CGCGCGATTT TGCCCAAAGT
 801 GAGCAAACCT TATGCTACTT ACGGTTTGGA CGATACCGCC GACATCTACG
 851 CCACCGACAT CGAAAACGTC GGCGCGCAAA TGAAATTCAC CGTCCATGTT
 901 CAAATGAAAG GACATGAGCA GGGGTCGTTT GAAGTCGTGC TGAATATGCC
 951 CGGCAGACAC AACGTGCTGA ACGCATTGGC AGCCATCGGC gtggcGCTgg
1001 aagtcGgCGC ATcggttgAA GCGAtcCAAA AaggCTTGCT CGGCTTTGAA
1051 GGCGTCGGCC GCCGCTTCCA AAAATAcggc gacatCAagt tgccaaacgg
     cggGaccgCT TTgctGGTGG ACGATTAcgg ACACCACCCC GTCGAAATGG
1151
      CGGcaaccct tgccgcTGCA CGCGGCGCGT ATCCGGAAAA acgtTTGGTG
     CtcgCCTTCC AGCCGCACCG CTATACCCGC ACGCGCGATT TGTTTGAAGA
      CTTTACCAAA GTACTCAATA CCGTTGatgC GCTGGTACTG ACCGAAGTTT
1251
     AtgccgccgG CGAAGAGCCG GTTGCCGCCG CCGactcCCG CGCCTTGGCG
1301
     CGTGCTATCC GCGTATTGGG CAAACTTGAG CCGATTTACT GCGAAAatgt
1401 CGCCGACCTG CCGCAAATGC TGATGAATGT TTTACAGGAT GGCGatgttg
     tgttgAATAT GggTgcggga agcatcaacc gcgttccttc cgcgctgttg
1501 gaattgtcga AACAGAtttg A
```

This corresponds to the amino acid sequence <SEQ ID 310; ORF 092.ng>:

1	MFFISIRYIF	VRKLWCANGQ	TFKITPLRTK	NOPERNIMME	MOMENTHEMA
51	IGGVGMSGIA	EVLHNLGFKV	SGSDOARNAA	TEHLSSIGIO	WADGREEN
101	NGADVVVAST	AVKKENPEVV	AALEROIPVI	PRALMIARIM	DEDUCTATAC
151	THGKTTTTSL	TASILGAAGL	DPTFVIGGKI.	NAAGTMARIG	RERUGIATAG
201	ESDASFLHLT	PIMSVVTNID	EDHMDTYGHS	VEKLHOAFID	RELIVATAD

289

```
251 AFLCVDSEHV RAILPKVSKP YATYGLDDTA DIYATDIENV GAQMKFTVHV
          301 QMKGHEQGSF EVVLNMPGRH NVLNALAAIG VALEVGASVE AIQKGLLGFE
               GVGRRFQKYG DIKLPNGGTA LLVDDYGHHP VEMAATLAAA RGAYPEKRLV
          401 LAFQPHRYTR TRDLFEDFTK VLNTVDALVL TEVYAAGEEP VAAADSRALA
          451 RAIRVLGKLE PIYCENVADL PQMLMNVLQD GDVVLNMGAG SINRVPSALL
          501 ELSKQI*
The following partial DNA sequence was identified in N. meningitidis <SEO ID 311>:
     m092.seg
               ATGTTTTTA TTTCAATCCG CTATATATTT GTCAGAAAAC TATGGCGCGC
            1
              AAACGGTCAG CCCTTTAAAA TAACGCCTTT ACGCATCGAA AATCCACCGG
           51
              AACGCAACAT TATGATGAAA AATCGAGTTA CCAACATCCA TTTTGTCGGT
          101
          151
               ATCGGCGGCG TCGGCATGAG CGGCATCGCC GAAGTCTTGC ACAATTTGGG
               CTTTAAAGTT TCCGGTTCGG ATCAGGCGCG AAATGCCGCT ACCGAGCATT
          201
          251 TGGGCAGCCT GGGCATTCAA GTTTATCCCG GCCATACCGC CGAACACGTT
          301 AACGGTGCGG ATGTCGTCGT TACCTCTACC GCCGTCAAAA AAGAAAATCC
          351 CGAAGTTGTC GCTGCGTTGG AGCAGCAAAT TCCCGTTATT CCGCGCGCCC
          401 TGATGTTGGC GGAGTTGATG CGCTTCCGTG ACGGCATCGC CATTGCCGGC
          451 ACGCACGGCA AAACCACGAC CACCAGCCTG ACCGCCTCCA TCCTCGGCGC
          501 GGCAGGACTT GACCCGACTT TCGTTATCGG CGGCAAACTC AACGCCGCAG
          551 GCACTAACGC CCGCTTGGGC AAAGGCGAAT ACATCGTTGC CGAAGCCGAC
          601 GAGTCGGATG CATCCTTTCT GCACCTGACA CCGATTATGT CCGTCGTTAC
          651 CAATATCGAC GAAGACCATA TGGATACCTA CGGGCACAGC GTCGAAAAAC
          701 TGCATCAGGC GTTTATCGAT TTCATCCACC GTATGCCCTT CTACGGCAAA
          751 GCCTTTTGT GTATTGACAG CGAACACGTC CGCGCGATTT TGCCCAAAGT
          801 GAGCAAACCT TATGCTACTT ACGGTTTGGA CGATACCGCC GACATCTACG
          851
              CCACCGACAT CGAAAACGTC GGCGCGCAAA TGAAATTCAC CGTCCATGTT
               CAAATGAAAG GACATGAGCA GGGGTCGTTT GAAGTCGTGC TGAATATGCC
               CGGCAGACAC AACGTGCTGA ACGCATTGGC AGCCATCGGC GTGGCGCTGG
         1001 AAGTCGGCGC ATCGGTTGAA GCGATCCAAA AAGGCTTGCT CGGCTTTGAA
         1051 GGCGTCGGCC GCCGCTTCCA AAAATACGGC GACATCAAGT TGCCAAACGG
         1101 CGGGACCGCG CTCTTGGTGG ACGACTACGG ACACCACCCC GTCGAAATGG
         1151 CGGCGACCCT TGCCGCCGCA CGCGGCGCGT ATCTGGAAAA ACGTTTGGTA
         1201 CTCGCCTTCC AGCCGCACCG CTATACCCGC ACGCGCGATT TGTTTGAAGA
         1251 CTTTACCAAA GTCCTCAATA CCGTTGACGC GCTGGTGCTG ACCGAAGTTT
         1301 ATGCCGCCGG TGAAGAGCCG ATTGCCGCCG CCGATTCCCG CGCTCTTGCC
         1351 CGCGCCATCC GCGTGTTGGG CAAACTCGAG CCGATTTACT GCGAAAACGT
         1401 TGCCGATCTG CCCGAAATGC TGTTGAACGT TTTGCAGGAC GGCGACATCG
         1451 TGTTGAATAT GGGCGCGGGA AGCATCAACC GCGTCCCCGC CGCGCTGCTG
               GCATTGTCGA AACAGATTTG A
This corresponds to the amino acid sequence <SEQ ID 312; ORF 092>:
     m092.pep
               MFFISIRYIF VRKLWRANGO PFKITPLRIE NPPERNIMMK NRVTNIHFVG
           51 IGGVGMSGIA EVLHNLGFKV SGSDQARNAA TEHLGSLGIQ VYPGHTAEHV
          101 NGADVVVTST AVKKENPEVV AALEQQIPVI PRALMLAELM RFRDGIAIAG
          151 THGKTTTTSL TASILGAAGL DPTFVIGGKL NAAGTNARLG KGEYIVAEAD
          201 ESDASFLHLT PIMSVVTNID EDHMDTYGHS VEKLHQAFID FIHRMPFYGK
          251 AFLCIDSEHV RAILPKVSKP YATYGLDDTA DIYATDIENV GAQMKFTVHV
               QMKGHEQGSF EVVLNMPGRH NVLNALAAIG VALEVGASVE AIQKGLLGFE
               GVGRRFQKYG DIKLPNGGTA LLVDDYGHHP VEMAATLAAA RGAYLEKRLV
          401 LAFQPHRYTR TRDLFEDFTK VLNTVDALVL TEVYAAGEEP IAAADSRALA
               RAIRVLGKLE PIYCENVADL PEMLLNVLQD GDIVLNMGAG SINRVPAALL
          451
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae

ORF 092 shows 96.6% identity over a 506 aa overlap with a predicted ORF (ORF 092.ng) from N. gonorrhoeae:

m092/g092

501 ALSKQI*

10 20 30 40 50 60 m092.pep MFFISIRYIFVRKLWRANGQPFKITPLRIENPPERNIMMKNRVTNIHFVGIGGVGMSGIA

g092		 WCANGQTFKI 20	:     TPLRTKNQPE 30	: RNIMMKNRVS 40	  SNIHFVGIGG   50	 VGMSGIA 60
m092.pep	70 EVLHNLGFKVSGSD            EVLHNLGFKVSGSD 70	111111111	:		111:1111	
m092.pep g092	130 AALEQQIPVIPRALI     :        AALERQIPVIPRALI 130	111111	111111111			HIIIII
m092.pep	190 NAAGTNARLGKGEY:            NAAGTNARLGKGEY: 190			1111111111		
m092.pep	250 FIHRMPFYGKAFLCI	:			HHHHHH	
m092.pep g092	310 QMKGHEQGSFEVVLN            QMKGHEQGSFEVVLN 310				111111111	111111
m092.pep g092	370 DIKLPNGGTALLVDD	11111111				111111
m092.pep	430 VLNTVDALVLTEVYA           VLNTVDALVLTEVYA 430	1111:111	450 DSRALARAIRV          OSRALARAIRV 450			: 11111
m092.pep	490 GDIVLNMGAGSINRV   :          GDVVLNMGAGSINRV 490	1:111 111	11			

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 313>: a092.seq

1	ATGTTTTTTA	TTTCAATCCG	CTATATATTT	GTCAGAAAAC	TATGGCGCGC
51	AAACGGTCAG	CCCTTTAAAA	TAACGCCTTT	ACGCATCGAA	AATCCACCGG
101	AACGCAACAT	TATGATGAAA	AATCGAGTGA	CCAACATCCA	TTTTGTCGGT
151	ATCGGCGGCG	TCGGCATGAG	CGGTATCGCC	GAAGTCTTGC	ACAATTTGGG
201	TTTTAAAGTT	TCCGGTTCGG	ATCAGGCGCG	AAATGCCGCT	ACCGAGCATT
251	TGGGCAGCCT				

301	AACGGTGCGG	ATGTCGTCGT	TACCTCTACC	GCCGTCAAAA	AAGAAAATCC			
351	CGAAGTTGTC	GCTGCGTTGG	AGCAGCAAAT	TCCCGTTATT	CCGCGCGCCC			
401	TGATGTTGGC	GGAGTTGATG	CGCTTCCGTG	ACGGCATCGC	CATTGCCGGC			
451	ACGCACGGCA	AAACCACGAC	CACCAGCCTG	ACCGCCTCCA	TCCTCGGCGC			
501	GGCAGGACTT	GACCCGACTT	TCGTTATCGG	CGGCAAACTC	AACGCCGCAG			
551	GCACCAACGC	CCGCTTGGGC	AAAGGCGAAT	ACATCGTTGC	CGAAGCCGAC			
601	GAGTCGGATG	CATCCTTTCT	GCACCTGACA	CCGATTATGT	CCGTCGTTAC			
651	CAATATCGAC	GAAGACCATA	TGGATACCTA	CGGGCACAGT	GTTGAGAAGC			
701	TGCATCAGGC	GTTTATCGAT	TTCATCCACC	GTATGCCCTT	CTACGGCAAA			
751	GCCTTTTTGT	GTATTGACAG	CGAACACGTC	CGCGCGATTT	TGCCCAAAGT			
801	GAGCAAACCT	TATGCTACTT	ACGGTTTGGA	CGATACCGCC	GACATCTACG			
851	CCACCGACAT	CGAAAACGTC	GGCGCGCAAA	TGAAATTCAC	CGTCCATGTT			
901	CAAATGAAAG	GACATGAGCA	GGGGTCGTTT	GAAGTCGTGC	TGAATATGCC			
951	CGGCAGACAC	AACGTGCTGA	ACGCATTGGC	AGCCATCGGC	GTGGCGCTGG			
1001	AAGTCGGCGC	ATCGGTTGAA	GCGATCCAAA	AAGGCTTGCT	CGGCTTTGAA			
1051	GGTGTCGGCC	GCCGCTTCCA	AAAATACGGC	GACATCAAGT	TGCCAAACGG			
1101	TGGAACCGCG	CTCTTGGTGG	ACGACTACGG	ACACCACCCC	GTCGAAATGG			
1151	CGGCGACCCT	TTCCGCCGCA	CGCGGCGCGT	ATCCGGAAAA	ACGTTTGGTA			
1201	CTCGCCTTCC	AGCCGCACCG	CTATACCCGC	ACGCGCGATT	TGTTTGAAGA			
1251	CTTTACCAAA	GTCCTCAATA	CCGTTGACGC	GCTGGTGCTG	ACCGAAGTTT			
1301	ATGCCGCCGG	TGAAGAGCCG	ATTGCCGCCG	CTGATTCCCG	CGCTCTTGCC			
1351	CGCGCCATCC	GCGTGTTGGG	CAAACTCGAG	CCGATTTACT	GCGAAAACGT			
1401	TGCCGATCTG	CCCGAAATGC	TGTTGAACGT	TTTGCAGGAC	GGCGACATCG			
1451	TGTTGAATAT	GGGTGCGGGA	AGCATCAACC	GCGTCCCCGC	CGCGCTGCTG			
1501	GAATTGTCGA	AACAGATTTG	A					
s corre	s corresponds to the amino acid sequence <seo 092.a="" 314;="" id="" orf=""></seo>							

# This corresponds to the amino acid sequence <SEQ ID 314; ORF 092.a>: a092.pep

1	MFFISIRYIF	VRKLWRANGQ	PFKITPLRIE	NPPERNIMMK	NRVTNIHFVG
51	IGGVGMSGIA	EVLHNLGFKV	SGSDQARNAA	TEHLGSLGIQ	VYPGHTAEHV
101	NGADVVVTST	AVKKENPEVV	AALEQQIPVI	PRALMLAELM	RFRDGIAIAG
151	THGKTTTTSL	TASILGAAGL	DPTFVIGGKL	NAAGTNARLG	KGEYIVAEAD
201	ESDASFLHLT	PIMSVVTNID	EDHMDTYGHS	VEKLHQAFID	FIHRMPFYGK
251	AFLCIDSEHV	RAILPKVSKP	YATYGLDDTA	DIYATDIENV	GAOMKFTVHV
301	QMKGHEQGSF	EVVLNMPGRH	NVLNALAAIG	VALEVGASVE	AIQKGLLGFE
351	GVGRRFQKYG	DIKLPNGGTA	LLVDDYGHHP	VEMAATLSAA	RGAYPEKRLV
401		TRDLFEDFTK			
451	RAIRVLGKLE	PIYCENVADL	PEMLLNVLQD	GDIVLNMGAG	SINRVPAALL
501	ELSKQI*				

m092/a092 99.4% identity over a 506 aa overlap

	10	20	30	40	50	60
m092.pep	MFFISIRYIFVRKLWR	ANGQPFKITP	LRIENPPERN	IMMKNRVTNI	HFVGIGGVGM.	SGIA
			1111111111	1111111111		
a092	MFFISIRYIFVRKLWR	ANGQPFKITP	LRIENPPERN	IMMKNRVTNI	HFVGIGGVGM.	SGIA
	10	20	30	40	50	60
	70	80	90	100	110	120
m092.pep	EVLHNLGFKVSGSDQA	RNAATEHLGS	LGIOVYPGHT	'AEHVNGADVV	VTSTAVKKEN	PEVV
						1111
a092	EVLHNLGFKVSGSDQA	· · · · · · · · · · · ·		AFHVNGADVV	UTSTAUKKEN:	1 1 1 1 DE 1717
	70	80	90	100	110	120
	. 0	00	50	100	110	120
	130	140	150	160	170	180
m092.pep	AALEQQIPVIPRALML					
mosz.pep				TISLIASILG	AAGLUPTEVI	3GKL
a092				!	11111111	1111
a092	AALEQQIPVIPRALML					
	130	140	150	160	170	180
	190	200	210	220	230	240
m092.pep	NAAGTNARLGKGEYIV	AEADESDASF:	LHLTPIMSVV	TNIDEDHMDT	YGHSVEKLHQ	AFID
			1111111111	1111111111	111111111	
a092	NAAGTNARLGKGEYIV	AEADESDASF:	LHLTPIMSVV	TNIDEDHMOT	YCHSVEKTHO	חדת

, T.

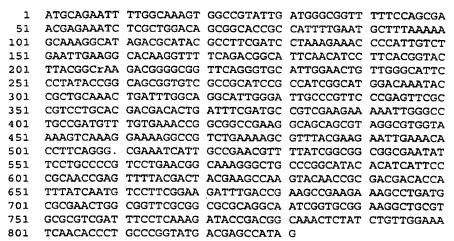
	190	200	210	220	230	240
m092.pep	250 FIHRMPFYGKAFLC            FIHRMPFYGKAFLC 250	1       1   1   1   1   1		 YGLDDTADIY	 ATDIENVGAQ	 MKFTVHV
m092.pep	310 QMKGHEQGSFEVVLI	320 MPGRHNVLN	330 ALAAIGVAL	280 340 EVGASVEAIQF	290 350 (GLLGFEGVG)	300 360 RRFQKYG
a092					TILLIA LIBI	
m092.pep	370 DIKLPNGGTALLVDI            DIKLPNGGTALLVDI 370		111:11111		I I I I I I I I I I I I I I I I I I I	 LFEDFTK
m092.pep	430 VLNTVDALVLTEVYA             VLNTVDALVLTEVYA 430		450 DSRALARAIR	460 VLGKLEPIYC	470 ENVADLPEMI	
m092.pep	490 GDIVLNMGAGSINRV            GDIVLNMGAGSINRV 490	11111 1111	111			

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 315>:

```
aTGCAGAATt ttgGCAAAGT ggccgtATTG ATGGGtggtT TTTCCAGCGA
 51 ACGAGAAAtc tcgcTGGACA GCGGTACCGC CATTTTGAAC GCCTTAAAAA
     GCAAAGGCAT AGACGCATAC GCCTTCGACC CTAAGGAAAC GCCGTTATCC
151 GAACTGAAGG AGCGGGGCTT TCAGACGGCA TTCAACATCC TTCACGGTAC
201 TTACGGCGAA GACGGGGCTG TTCAGGGTGC ATTGGAACTG TTGGGCATTC
251 CCTATACCGG CAGCGGTGTC GCCGCCTCCG CCATCGGCAT GGACAAATAC
301 CGCTGCAAAC TGATTTGGCA GGCATTGGGA TTACCCGTTC CCGAGTTCGC
351 CGTACTGTAC GATGATACCG ATTTCGATGC CGTCGAAGAA AAATTGGGTC
401 TGCCGATGTT TGTGAAGCCG GCGGCCGAAG GCAGCAGCgt cggcgtggta
451 aAAGTCAAAG AAAaaggccg TCTGAAAAGC GTTtacgaag aatTGAaaCA
501 CCTTcagggg cgaAAtcatt gccgAacgTT TTATCGGCGG CGGCGAATAT
    TCCTGCCCCG TCCTGAACGG CAAAGGGCTG CCCGGCATAC ACATCATCCC
    CGCAACCGAG TTTTACGACt acgaagccaa GtacaaCCGA GACGAcacca
651 tttaTCAATG TCCTTCGGAA GATTTGACCG AAGCCGAAGA AAGCCTGATG
701 CGCGAACTGG CGGTTCGCGG CGCACAGGCA ATCGGTGCGG AAGGCTGCGT
751 GCGCGTCGAT TTCCTCAAAG ATACCGACGG CAAACTCTAT CTGTTGGAAA
    TCAACACCCT GCCCGGTATG ACCGGCCATA G
```

This corresponds to the amino acid sequence <SEQ ID 316; ORF 093.ng>:

- 1 MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPLS
  51 ELKERGFQTA FNILHGTYGE DGAVQGALEL LGIPYTGSGV AASAIGMDKY
  101 RCKLIWQALG LPVPEFAVLY DDTDFDAVEE KLGLPMFVKP AAEGSSVGVV
  151 KVKEKGRLKS VYEELKHLQG RNHCRTFYRR RRIFLPRPER QRAARHTHHP
  201 RNRVLRLRSQ VQPRRHHLSM SFGRFDRSRR KPDARTGGSR RTGNRCGRLR
  251 ARRFPQRYRR QTLSVGNQHP ARYDRP*
- The following partial DNA sequence was identified in N. meningitidis <SEQ ID 317>:



This corresponds to the amino acid sequence <SEQ ID 318; ORF 093>: m093.pep

- MONFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPLS 51 ELKAQGFQTA FNILHGTYGX DGAVQGALEL LGIPYTGSGV AASAIGMDKY
- 101 RCKLIWQALG LPVPEFAVLH DDTDFDAVEE KLGLPMFVKP AAEGSSVGVV KVKGKGRLKS VYEELKHLQX RNHCRTFYRR RRIFLPRPER QRAARHTHHS
- RNRVLRLRSQ VOPRRHHLSM SFGRFDRSRR KPDARTGGSR RAGNRCGRLR
- ARRFPORYRR QTLSVGNQHP ARYDEP*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae

ORF 093 shows 96.7% identity over a 276 aa overlap with a predicted ORF (ORF 093.ng) from N. gonorrhoeae:

m093/g093

	10	20	30	40	50	60
m093.pep	MQNFGKVAVLMGGF	SSEREISLDS	SGTAILNALKS	KGIDAYAFDI	KETPLSELKA	QGFQTA
						: [ ] [ ] [
g093	MONFGKVAVLMGGF	SSEREISLDS	SGTAILNALKS	KGIDAYAFDI	KETPLSELKE	RGFOTA
	10	20	30	40	50	60
	70	80	90	100	110	120
m093.pep	FNILHGTYGXDGAV	QGALELLGI	PYTGSGVAASA	IGMDKYRCKI	JIWOALGLPVE	EFAVLH
					111111111	11111:
g093	FNILHGTYGEDGAV	OGALELLGI1	PYTGSGVAASA	IGMDKYRCKI	JWOALGLPVE	
	70	80	90	100	110	120
	130	140	150	160	170	180
m093.pep	DDTDFDAVEEKLGL	PMFVKPAAE	GSSVGVVKVKG	KGRLKSVYE	LKHLOXRNHO	RTFYRR
				1111111111		
g093	DDTDFDAVEEKLGL	PMFVKPAAE	SSVGVVKVKE	KGRLKSVYE	LKHLOGRNHO	RTFYRR
	130	140	150	160	170	180
	190	200	210	220	230	240
m093.pep	RRIFLPRPERQRAA	RHTHHSRNRY	VLRLRSQVQPR	RHHLSMSFGF	REDRSRRKPDA	
		11111 111		1111111111	1111111111	
g093	RRIFLPRPERORAA	RHTHHPRNRY	VLRLRSOVOPR	RHHLSMSFGF	FDRSRRKPDA	RTGGSR
	190	200	210	220	230	240
						2.0
	250	260	270			
m093.pep	RAGNRCGRLRARRF	PORYRROTLS	SVGNOHPARYD	EPX		
				: [ ]		
g093	RTGNRCGRLRARRF	PORYRROTLS				
	250	260	270			

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 319>: a093.seq

1	ATGCAGAATT	TTGGCAAAGT	GGCCGTATTG	ATGGGCGGTT	TTTCCAGCGA
51	ACGAGAAATC	TCGCTGGACA	GCGGCACCGC		
101	GCAAAGGCAT	AGACGCATAC	GCCTTCGATC		COLLINERAN
151	GAATTGAAGG	CACAAGGTTT		001 11 10 01 11 11 10	00011110101
201	TTACGGCGAA	GACGGGGCTG			
251	CCTATACCGG	CAGCGGTGTC			TTGGGCATTC
301		TGATTTGGCA			GGACAAATAC
351	CGTCCTGCAC		ATTTCGATGC		CCGAGTTCGC
401	TGCCGATGTT	TGTGAAACCG		CGTCGAAGAA	
451	AAAGTCAAAG		00000001110	GCAGCAGCGT	
501	CTTTCAGGG.	9	TCTGAAAAGC	GTTTACGAAG	
551	TCCTGCCCTG	CGAAATCATT	GCCGAACGGT	TTATCGGCGG	
601		TGTTGAACGG	CAAAGGCCTG	CCCGGCATAC	ACATCATCCC
	CGCGACCGAG		ACGAAGCCAA	GTACAACCGC	AACGACACCA
651	TTTATCAATG	TCCTTCGGAA	GATCTGACCG	AAGCCGAAGA	
701	CGCGAACTGG	CGGTTCGCGG	CGCGCAGGCA		AAGGCTGCGT
751	GCGCGTCGAT	TTCCTCAAAG		CAAACTCTAT	
801	TCAACACCCT	GCCCGGTATG	ACCGGCCATA	G	

## This corresponds to the amino acid sequence <SEQ ID 320; ORF 093.a>: a093.pep

- 1 MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPLS 51 ELKAQGFQTA FNILHGTYGE DGAVQGALEL LGIPYTGSGV AASAIGMDKY 101 RCKLIWQALG LPVPEFAVLH DDTDFDAVEE KLGLPMFVKP AAEGSSVGVV
- 151 KVKGKGRLKS VYEELKHFQX RNHCRTVYRR RRIFLPCVER QRPARHTHHP 201 RDRVL*LRSQ VQPQRHHLSM SFGRSDRSRR KPDARTGGSR RAGNRCGRLR
- 251 ARRFPQRYRR QTLSVGNQHP ARYDRP*

#### 95.7% identity over a 276 aa overlap m093/a093

250

260

m093.pep	10 MQNFGKVAVLMGG            MQNFGKVAVLMGG: 10	1111111111	111111111	11111111111		
m093.pep	70 FNILHGTYGXDGAV           FNILHGTYGEDGAV 70			11111111111	11111111	120 PEFAVLH
m093.pep	130 DDTDFDAVEEKLGI           DDTDFDAVEEKLGI 130		111111111	1111111111	Add to the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of the control of	
m093.pep	190 RRIFLPRPERQRAA          !     RRIFLPCVERQRPA 190	7   1   1   1   5   1	1 111111			
m093.pep	250 RAGNRCGRLRARRF           RAGNRCGRLRARRF 250	11111111		• 1 1		

270

```
The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 321>:
     q094.seq
            1
               ATGTATTCGC CTTTGCCCAA GCGGGCGTTG GTGCCTGCGG CGTTGAGTTT
           51 GCCGCCGATA ACGAAAGTGG GGTCGAGTCC TGCCGCGCCG AGGATGGAGG
          101 CGGTCAGGCT GGTGGTCGTG GTTTTGCCGT GCGTAccggc aatggcgatg
          151 CCGTCACGGA AGCGCATCAG CTCTGCCAGC ATCAAGGCGC GCGGAATAAC
          201 GGGAATTTGC CGCTCCAACG CAgcgacaAC TTCGGgattT TCTTTCTTGA
               CGGCGGTAGA GGCAACGACG ACATCCGCAC CGTTAACGTG TTCTGCGGTA
               TGGCCGGGAT AA
This corresponds to the amino acid sequence <SEQ ID 322; ORF 094.ng>:
     g094.pep
               MYSPLPKRAL VPAALSLPPI TKVGSSPAAP RMEAVRLVVV VLPCVPAMAM
               PSRKRISSAS IKARGITGIC RSNAATTSGF SFLTAVEATT TSAPLTCSAV
           51
          101
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 323>:
     m094.seq
               ATGTATTCGC CTTTGCCCAA GCGGGCGTTA GTGCCTGCGG CGTTGAGTTT
           51 GCCGCCGATA ACGAAAGTCG GGTCAAGTCC TGCCGCGCCG AGGATGGAGG
          101 CGGTCAGGCT GGTGGTCGTG GTTTTGCCGT GCGTGCCGGC AATGGCGATG
          151 CCGTCACGGA AGCGCATCAA CTCCGCCAAC ATCAGGGCGC GCGGAATAAC
          201 GGGAATTTGC TGCTCCAACG CAGCGACAAC TTCGGGATTT TCTTTTTGA
          251 CGGCGGTAGA GGTAACGACG ACATCCGCAC CGTTAACGTG TTCGGCGGTA
              TGGCCGGGAT AA
This corresponds to the amino acid sequence <SEQ ID 324; ORF 094>:
     m094.pep
               MYSPLPKRAL VPAALSLPPI TKVGSSPAAP RMEAVRLVVV VLPCVPAMAM
               PSRKRINSAN IRARGITGIC CSNAATTSGF SFLTAVEVTT TSAPLTCSAV
           51
          101
              WPG*
Computer analysis of this amino acid sequence gave the following results:
Homology with a predicted ORF from N. gonorrhoeae
ORF 094 shows 95.1% identity over a 103 as overlap with a predicted ORF (ORF 094.ng)
from N. gonorrhoeae:
     m094/q094
                                   20
     m094.pep
                  MYSPLPKRALVPAALSLPPITKVGSSPAAPRMEAVRLVVVVLPCVPAMAMPSRKRINSAN
                  q094
                  MYSPLPKRALVPAALSLPPITKVGSSPAAPRMEAVRLVVVVLPCVPAMAMPSRKRISSAS
                                             30
                                                      40
                                                                50
                         70
                                   80
                                             90
                  IRARGITGICCSNAATTSGFSFLTAVEVTTTSAPLTCSAVWPGX
     m094.pep
                  g094
                  IKARGITGICRSNAATTSGFSFLTAVEATTTSAPLTCSAVWPGX
                         70
                                   80
                                             90
                                                      100
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 325>:
a094.seg
      1
         ATGTATTCGC CTTTGCCCAA GCGGGCGTTG GTGCCTGCGG CGTTGAGTTT
         GCCGCCGATA ACGAAAGTCG GGTCAAGTCC TGCCGCGCCG AGGATGGAGG
     51
    101 CGGTCAGGCT GGTGGTCGTG GTTTTGCCGT GCGTGCCGGC AATGGCGATG
    151 CCGTCACGGA AGCGCATCAA CTCCGCCAAC ATCAGGGCGC GCGGAATAAC
    201 GGGAATTTGC TGCTCCAACG CAGCGACAAC TTCGGGATTT TCTTTTTGA
         CGGCGGTAGA GGTAACGACG ACATCCGCAC CGTTAACGTG TTCTGCGGTA
    251
    301
         TGGCCGGGAT AA
```

This corresponds to the amino acid sequence <SEQ ID 326; ORF 094.a>:

a094.pep MYSPLPKRAL VPAALSLPPI TKVGSSPAAP RMEAVRLVVV VLPCVPAMAM 1 PSRKRINSAN IRARGITGIC CSNAATTSGF SFLTAVEVTT TSAPLTCSAV 51 101 WPG* 100.0% identity over a 103 aa overlap m094/a09410 20 30 40 50 MYSPLPKRALVPAALSLPPITKVGSSPAAPRMEAVRLVVVVLPCVPAMAMPSRKRINSAN m094.pep MYSPLPKRALVPAALSLPPITKVGSSPAAPRMEAVRLVVVVLPCVPAMAMPSRKRINSAN a094 20 30 40 50 10 70 80 90 100 IRARGITGICCSNAATTSGFSFLTAVEVTTTSAPLTCSAVWPGX m094.pep IRARGITGICCSNAATTSGFSFLTAVEVTTTSAPLTCSAVWPGX a094 70 80 The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 327>: q095.seq ATGTCCTTTC ATTTGAACAT GGACGGTGAA TTTCATTTGC GCGCCGACGT 1 51 TTTCGATGTC GGTGGCGTAG ATGTCGGCGG TATCGTCCAA ACCGTAAGTA 101 GCATAAGGTT TGCTCACTTT GGGCAAAATC GCGCGGACGT GTTCGCTGTC 151 AACACACAA AAGGCTTTGC CGTAGAAGGG CATACGGTGG ATGAAATCGA 201 TAAACGCCTG ATGCAGTTTT TCGACGCTGT GCCCGTAGGT ATCCATATGG 251 TCTTCGTCGA TATTGGTAAC GACGGACATA ATCGGGGTCA GTGTAGGAAA 301 GAGGCATCGG ATCGTCGGCT TCGGCAACGA TGTATTCGCC TTTGCCCAAG CGGGCGTTGG TGCCTGCGGC GTTGA 351 This corresponds to the amino acid sequence <SEQ ID 328; ORF 095.ng>: g095.pep MSFHLNMDGE FHLRADVFDV GGVDVGGIVQ TVSSIRFAHF GQNRADVFAV NTOKGFAVEG HTVDEIDKRL MQFFDAVPVG IHMVFVDIGN DGHNRGQCRK 51 EASDRRLRQR CIRLCPSGRW CLRR* The following partial DNA sequence was identified in N. meningitidis <SEQ ID 329>: m095.seq ATGTCCTTTC ATTTGAACAT GGACGGTGAA TTTCATTTGC GCGCCGACGT TTTCGATGTC GGTGGCGTAG ATGTCGGCGG TATCGTCCAA ACCGTAAGTA 51 101 GCATAAGGTT TGCTCACTTT GGGCAAAATC GCGCGGACGT GTTCGCTGTC 151 AATACACAAA AAGGCTTTGC CGTAGAAGGG CATACGGTGG ATGAAATCGA 201 TAAACGCCTG ATGCAGTTTT TCGACGCTGT GCCCGTAGGT ATCCATATGG 251 TCTTCGTCGA TATTGGTAAC GACGGACATA ATCGGTGTCA GTGCAGAAAG GATGCATCCG ACCGTCGGCT TCGGCAACGA TGTATTCGCC TTTGCCCAAG CGGGCGTTAG TGCCTGCGGC GTTGA This corresponds to the amino acid sequence <SEQ ID 330; ORF 095>: m095.pep MSFHLNMDGE FHLRADVFDV GGVDVGGIVQ TVSSIRFAHF GQNRADVFAV NTOKGFAVEG HTVDEIDKRL MOFFDAVPVG IHMVFVDIGN DGHNRCQCRK DASDRRLROR CIRLCPSGRX CLRR* Computer analysis of this amino acid sequence gave the following results: Homology with a predicted ORF from N. gonorrhoeae ORF 095 shows 97.6% identity over a 124 aa overlap with a predicted ORF (ORF 095.ng) from *N. gonorrhoeae*: m095/g095 MSFHLNMDGEFHLRADVFDVGGVDVGGIVQTVSSIRFAHFGQNRADVFAVNTQKGFAVEG m095.pep 

297

g095	MSFHLN	MDGEFHLRAD	VFDVGGVD	VGGIVQTVSS	IRFAHFGONR	ADVFAVNTQI	(GFAVEG	
J		10	20	30	40	50	60	
		70	80	90	100	110	120	
m095.g		DKRLMQFFDA						
g095								
9093	111 4 2 2 3	70	80	90	100	110	120	
m095. <b>r</b>								
	11111							
g095	CLRRX							
C 11 :	1 5514			C - 1 : M		«CEO ID	2015	
	ng partial DNA	sequence v	was identi	nea in N. n	neningitiais	<2EQ ID	331>:	
a095.s	•							
		C ATTTGAAC C GGTGGCGT						
,		T TGCTCACT						
		A AAGGCTTT						
		G ATGCAGCT						
		A TATTGGTA						
3		G ACCGTCGG						
3	S51 CGGGCGTTG	G TGCCTGCG	GC GTTGA					
This corresp	onds to the am	ino acid sed	quence <s< td=""><td><b>EQ ID 332</b></td><td>2; ORF 095.</td><td>.a&gt;:</td><td></td><td></td></s<>	<b>EQ ID 332</b>	2; ORF 095.	.a>:		
a095.p			-	_				
•		E FHLRADVE	DV GGVDV	GGIVQ TVSS	IRFAHF GQN	RADVFAV	-	
		G HTVDEIDK		rvpvg IHMV	FVDIGN DGH	NRCQCRK		
1	01 DASDRRLRÇ	R CIRLCPSG	RW CLRR*					
225 /	00.00		- 104					
m095/a	1095 96.0%	identity i	.n 124 aa	overlap				
		10	20	30	40	50	60	
m095.p	nep MSFHLN	MDGEFHLRAD						• •
•		1111111111						30
a095	MSFHLN	MDGEFHLRAD						
		10	20	30	40	50	60	
		70	80	90	100	110	120	
m095.g	oo UTVDEI	7U DKRLMQFFDA						
11093.1								
a095		DKRLMOLLNI						
		70	80	90	100	110	120	
m095.p	•							
a095	 CLRRX							
8093	CHILIA							

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID

333>:

```
g096.seq1ATGGCCGGTCATACCGGGCAGGGTGTTGATTTCCAACAGATAGAGTTTGC51CGTCGGTATCTTTGAGGAAATCGACGCGCACGCAGCCTTCCGCACCGATT101GCCTGTGCGCCGCGAACCGCCAGTTCGCGCATCAGGCTTTCTTCGGCTTC151GGTCAAATCTTCCGAAGGACATTGAtaaatggtgTCGTCTCGGttgtaCt
```

```
201 tggcttcgta gTCGTAAAAC TCGGTTGCGG GGATGATGTG TATGCCGGGC
251 AGCCCTTTGC CGTTCAGGAC GGGGCAGGAA TATTCGCCGC CGCCGATAAA
301 AcgtTcggca atgaTTtcgc ccctgAAGGT GttTCAattc ttcgtaAACG
351 CTTTTCAGAc ggccttTTTC TTTGA
```

#### This corresponds to the amino acid sequence <SEQ ID 334; ORF 096.ng>: g096.pep

- MAGHTGQGVD FQQIEFAVGI FEEIDAHAAF RTDCLCAANR QFAHQAFFGF
- GQIFRRTLIN GVVSVVLGFV VVKLGCGDDV YAGQPFAVQD GAGIFAAADK
- 101 TFGNDFAPEG VSILRKRFSD GLFL*

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 335>:

```
m096.seq
          ATGGCTCGTC ATACCGGGCA GGGTGTTGAT TTCCAACAGA TAGAGTTTGC
         CGTCGGTATC TTTGAGGAAA TCGACGCGCA CGCAGCCTTC CGCACCGATT
      51
          GCCTGCGCGC CGCGAACCGC CAGTTCGCGC ATCAGGCTTT CTTCGGCTTC
     101
     151 GGTCAAATCT TCCGAAGGAC ATTGATAAAT GGTGTCGTCG CGGTTGTACT
     201 TGGCTTCGTA GTCGTAAAAC TCGGTTGCGG GAATGATGTG TATGCCGGGC
     251 AGCCCTTTGC CGTTCAGGAC GGGGCAGGAA TATTCGCCGC CGCCGATAAA
     301 ACGTTCGGCA ATGATTTCGC CC.TGAAGGT GTTTCAATTC TTCGTAAACG
351 CTTTTCAGAC GGCCTTTTCC TTTGA
```

#### This corresponds to the amino acid sequence <SEQ ID 336; ORF 096>: m096.pep

- MARHTGQGVD FQQIEFAVGI FEEIDAHAAF RTDCLRAANR QFAHQAFFGF 1 51
- GQIFRRTLIN GVVAVVLGFV VVKLGCGNDV YAGQPFAVQD GAGIFAAADK
- 101 TFGNDFAXEG VSILRKRFSD GLFL*

## m096/g096 96.0% identity in 124 aa overlap

m096.pep	10 MARHTGQGVDFQQI	20 EFAVGIFEE:	30 IDAHAAFRTDO	40 LRAANROFAH	50 IOAFFGFGOTE	60 MT 17797
g096	MAGHTGQGVDFQQI	EFAVGIFEE:	, , , , , , , , , , , ,	1 11111111		
	10	20	30	40	50	60
m096.pep	70 GVVAVVLGFVVVKL	80 GCGNDVYAGO	90 PFAVODGAGT	100	110	120
g096	:	1 1 1 5 1 1 1 1 1 1	- [	111111111	III IIIII	
	70	80	90	100	DFAPEGVSIL 110	RKRFSD 120

m096.pep GLFLX 11111 g096 GLFLX

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 337>:

#### a096.seg

```
1 ATGGCCGGTC ATACCGGGCA GGGTGTTGAT TTCCAACAGA TAGAGTTTGC
 51 CGTCGGTATC TTTGAGGAAA TCGACGCGCA CGCAGCCTTC CGCACCGATT
101 GCCTGCGCGC CGCGAACCGC CAGTTCGCGC ATCAGGCTTT CTTCGGCTTC
151 GGTCAGATCT TCCGAAGGAC ATTGATAAAT GGTGTCGTTG CGGTTGTACT
     TGGCTTCGTA GTCATAAAAC TCGGTCGCGG GGATGATGTG TATGCCGGGC
     AGGCCTTTGC CGTTCAACAC AGGGCAGGAA TATTCGCCGC CGCCGATAAA
251
301 CCGTTCGGCA ATGATTTCGC CCT.GAAAGT GTTTCAATTC TTCGTAAACG
351 CTTTTCAGAC GGCCTTTTCC TTTGA
```

## This corresponds to the amino acid sequence <SEQ ID 338; ORF 096.ng>: a096.pep

1 MAGHTGQGVD FQQIEFAVGI FEEIDAHAAF RTDCLRAANR QFAHQAFFGF

```
GQIFRRTLIN GVVAVVLGFV VIKLGRGDDV YAGQAFAVQH RAGIFAAADK
              PFGNDFAXES VSILRKRFSD GLFL*
          101
              92.7% identity in 124 aa overlap
m096/a096
                                   20
                                             30
                                                      40
                                                                50
                                                                          60
                 MARHTGQGVDFQQIEFAVGIFEEIDAHAAFRTDCLRAANRQFAHQAFFGFGQIFRRTLIN
     m096.pep
                  a096
                 MAGHTGQGVDFQQIEFAVGIFEEIDAHAAFRTDCLRAANRQFAHQAFFGFGQIFRRTLIN
                                                      40
                                   20
                         10
                                            30
                                                                50
                         70
                                   80
                                            90
                                                     100
                                                               110
                 GVVAVVLGFVVVKLGCGNDVYAGQPFAVQDGAGIFAAADKTFGNDFAXEGVSILRKRFSD
    m096.pep
                  GVVAVVLGFVVIKLGRGDDVYAGQAFAVQHRAGIFAAADKPFGNDFAXESVSILRKRFSD
     a096
                                   80
                                            90
                                                     100
                                                               110
                                                                        120
    m096.pep
                 GLFLX
                  a096
                 GLFLX
The following partial DNA sequence was identified in N. gonorrhoeae <SEO ID 339>:
     q097.seq
              ATGGATATTT CAAAACAAAC ATTGCTGGAT AGGGTTTTTA ACCTGAAGGC
           1
              AAACGGTACG ACGGTACGTA CCGAGTTGAT GGCGGGTTTG ACGACCTTTT
          51
              TGACGATGTG CTACATCGTT ATCGTCAATC CCCTGATTTT GGGCGAGACC
          101
         151 GGAATGGATA TGGGGGCGGT ATTCGTCGCT ACCTGTATCG CATCCGCCAT
         201 CGGCTGTTTT GTCATGGGTT TTATCGGCAA CTATCCGATT GCGCTTGCCC
              CGGGGATGGG GCTGAATGCC TATTTCACCT TTGCCGTCGT TAAGGGTATG
         251
         301 GGCGTGCCTT GGCAGGTGGC GTTGGGTGCG GTGTTCATTT CCGGTCTGAT
         351 TTTCATCCTG TTCAGCTTTT TTAAAGTCAG GGAAATGCTG GTCAACGCAC
         401 TGCCTATGGG TTTGAAAATG TCGATTGCCG CCGGTATCGG TTTGTTTTTG
         451 GCACTGATTT CCCTGAAAGG CGCAGGCATT ATCGTTGCCA ATCCGGCAAC
         501
              CTTGGTCGGC TTGGGCGATA TTCATCAGCC CAGCGCACTG TTGGCATTGT
              TCGGTTTTGT CATGGTGGTC GTATTGGGGT ATTTCCGCGT TCAAGGCGCA
         551
              ATCATCATCA CCATTCTGAC GATTACCGTC ATTGCCAGCC TGATGGGTTT
         601
              GAACGAGTTT CACGGCGTGG TCGGCGAAGT ACCGGGCATT GCGCCGACCT
         651
              TTATGCAGAT GGATTTTAAA GGTCTGTTTA CCGTCAGCAT GGTCAGCGTG
         701
         751 ATTTTCGTCT TCTTCTTGGT CGATTTGTTC GACAGTACCG GAACGCTGGT
         801 CGGCGTATCC CACCGTGCCG GACTGCTGGT GGACGGTAAG CTGCCCCGCC
         851
              TGAAACGCGC ACTGCTTGCA GACTCTACCG CCATTGTGGC AGGTGCGGCT
         901 TTGGGTACTT CTTCAACCAC GCCTTATGTG GAAAGCGCGG CGGGCGTATC
         951 GGCAGGCGGA CGGACCGGCC TGACGGCGGT TACCGTCGGC GTATTGATGC
        1001 TGGCGTGTCT GATGTTCTCC CCATTGGCGA AAAGTGTTCC GGTATTTGCC
        1051 ACCGCGCCG CACTGCTTTA TGTCGGCACG CAGATGCTCC GCAGTGCGAG
        1101 GGACATTGAT TGGGACGATA TGACTGAAGC CGCGCCCGCG TTCCTGACCA
              TTGTCTTCAT GCCGTTTACC TATTCGATTG CAGACGGCAT CGCCTTCGGC
        1151
        1201
              TTCATCAGCT ATGCCGTGGT CAAACTTTTG TGTCGCCGGA CTGGGGACGT
              GCCGCCTATG GTATGGGTTG TTGCCGTATT GTGGGCATTG AAATTCTGGT
        1251
        1301
              ATTTGGGCTG A
This corresponds to the amino acid sequence <SEQ ID 340; ORF 097.ng>:
    g097.pep
              MDISKQTLLD RVFNLKANGT TVRTELMAGL TTFLTMCYIV IVNPLILGET
           1
              GMDMGAVFVA TCIASAIGCF VMGFIGNYPI ALAPGMGLNA YFTFAVVKGM
          51
         101
              GVPWQVALGA VFISGLIFIL FSFFKVREML VNALPMGLKM SIAAGIGLFL
         151
              ALISLKGAGI IVANPATLVG LGDIHQPSAL LALFGFVMVV VLGYFRVQGA
              IIITILTITV IASLMGLNEF HGVVGEVPGI APTFMQMDFK GLFTVSMVSV
         201
              IFVFFLVDLF DSTGTLVGVS HRAGLLVDGK LPRLKRALLA DSTAIVAGAA
         251
         301
              LGTSSTTPYV ESAAGVSAGG RTGLTAVTVG VLMLACLMFS PLAKSVPVFA
         351 TAPALLYVGT QMLRSARDID WDDMTEAAPA FLTIVFMPFT YSIADGIAFG
```

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 341>:

```
ATGGACACTT CAAAACAAAC ACTGTTGGAC GGGATTTTTA AGCTGAAGGC
    1
       AAACGGTACK ACGGTGCGTA CCGAGTTGAT GGCGGGTTTG ACAACTTTTT
   51
       TGACGATGTG CTACATCGTT ATCGTCAACC CTCyGATTTT GGGCGAGACC
  101
       GGCATGGATA TGGGGGCGGT ATTCGTCGCT ACCTGTATCG CGTCTGCCAT
  201 CGGCTGTTTT GTTATGGGTT TTGTCGGCAA CTATCCGATT GCACTCGCAC
 251 CGGGGATGGG GCTGAATGCC TATTTCACCT TTGCCGTCGT TAAGGGTATG
 301 GGCGTGCCTT GGCAGGTTGC GTTGGGTGCG GTGTTCATCT CCGGTCTGAT
 351 TTTTATCCTG TTCAGCTTTT TTAAAGTCAG GGAAATGCTG GTCAACGCAC
 401 TGCCTATGGG TTTGAAAATG TCGATTGCTG CCGGTATCGG TTTGTTTTTG
 451 GCACTGATTT CCCTGAAAGG CGCAGGCATT ATCGTTGCCA ATCCGGCAAC
  501 CTTGGTCGGT TTGGGCGATA TTCATCAGCC GTCCGCGTTG TTGGCATTGT
 551 TCGGTTTTGC TATGGTGGTC GTATTGGGAC ATTTCCGCGT TCAAGGCGCA
      ATCATCATCA CCATCTTGAC CATTACCGTC ATTGCCAGCC TGATGGGTTT
      GAATGAATTT CACGGCATCA TCGGCGAAGT ACCGAGCATT GCGCCGACTT
      TTATGCAGAT GGATTTTGAA GGCCTGTTTA CCGTCAGCAT GGTCAGTGTG
 701
 751 ATTTTCGTCT TCTTCTTGGT CGATCTATTT GACAGTACCG GAACGCTGGT
 801 CGGCATATCC CACCGTGCCG GGCTGCTGGT GGACGGTAAG CTGCCCCGCC
 851 TGAAACGCGC ACTGCTTGCA GACTCTACCG CCATTGTGGC AGGTGCGGCT
 901 TTGGGTACTT CTTCCACCAC GCCTTATGTG GAAAGCGCGG CGGGCGTATC
 951 GGCAGGCGGA CGGACCGGCC TGACGGCGGT TACCGTCGGC GTATTGATGC
1001 TCGCCTGCCT GATGTTTTCA CCTTTGGCGA AAAGTGTTCC CGCTTTTGCC
1051 ACCGCGCCCG CCCTGCTTTA TGTCGGCACG CAGATGCTCC GCAGTGCGAG
1101 GGATATTGAT TGGGACGATA TGACGGAAGC CAGACCTCC GCAGTGCGAG
1151 TTGTTTCAT GCCGTTTACT TATTCGATTG CAGACGGAT CGCTTTCGGC
1201 TTCATCAGTT ATGCCGTGGT TAAACTTTTA TGCCGCCGA CGAAACACCT
      TTCATCAGTT ATGCCGTGGT TAAACTTTTA TGCCGCCGCA CCAAAGACGT
1251 TCCGCCTATG GTATGGATTG TTGCCGTATT GTGGGCACTG AAATTCTGGT
1301 ATTTGGGCTG A
```

This corresponds to the amino acid sequence <SEQ ID 342; ORF 097>:

MDTSKQTLLD GIFKLKANGT TVRTELMAGL TTFLTMCYIV IVNPXILGET
GMDMGAVFVA TCIASAIGCF VMGFVGNYPI ALAPGMGLNA YFTFAVVKGM
GVPWQVALGA VFISGLIFIL FSFFKVREML VNALPMGLKM SIAAGIGLFL
151 ALISLKGAGI IVANPATLVG LGDIHOPSAL LALFGFAMVV VLGHFRVQGA
201 IIITILTITV IASLMGLNEF HGIIGEVPSI APTFMQMDFE GLFTVSMVSV
251 IFVFFLVDLF DSTGTLVGIS HRAGLLVDGK LPRLKRALLA DSTAIVAGAA
301 LGTSSTTPYV ESAAGVSAGG RTGLTAVTVG VLMLACLMFS PLAKSVPAFA
351 TAPALLYVGT QMLRSARDID WDDMTEAAPA FLTIVFMPFT YSIADGIAFG

401 FISYAVVKLL CRRTKDVPPM <u>VWIVAVLWAL KFWYLG*</u>
Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae

ORF 097 shows 96.3% identity over a 436 aa overlap with a predicted ORF (ORF 097.ng) from N. gonorrhoeae:

m097/g097

	10	20	30	40	50	60
m097.pep	MDTSKQTLLDGIF	KLKANGTTVR'	TELMAGLTTF:	LTMCYIVIVN	PXILGETGMDN	ACTATURED IN
	11 [1][][] :]	:	111111111	11111111	11	111111
g097	MDISKQTLLDRVF	NLKANGTTVR'	TELMAGLTTF	LTMCYIVIVNI	PLILGETGMON	[[]]]]]]
	10	20	30	40	50	60
	70	80	90	100	110	120
m097.pep	TCIASAIGCFVMG	FVGNYPIALAI	PGMGLNAYFT	FAVVKGMGVPV	OVALGAVETS	בבט דרידוד זים!
g097	11111111			i	111111111	111111
_	TCIASAIGCFVMGI	TOMIFIALIA	GMGLNAY F.T.	AVVKGMGVPW	<i><b>IQVALGAVFIS</b></i>	GLIFIL
	70	80	90	100	110	120
	130	140	150	160	170	180

m097.pep	FSFFKVREMLVNALPM	(GLKMSIAAGI	GLFLALISLK	GAGIIVANPA	TLVGLGDIHC	
~0.07					TI VCI CD TVC	
g097	130	140	150	.GAGIIVANPA	170	180
	200				2.0	100
	190	200	210	220	230	240
m097.pep	LALFGFAMVVVLGHFF	RVQGAIIITIL	TITVIASLMG		VPSIAPTFMO	
			11111111	11111::11	11:11111	:
g097	LALFGFVMVVVLGYFF	-			_	
	190	200	210	220	230	240
	250	260	270	280	290	300
m097.pep	GLFTVSMVSVIFVFFI	VDLFDSTGTL	VGISHRAGLL	VDGKLPRLKR	ALLADSTAIV	AGAA
		ПППППП	11:111111	111111111		
g097		VDLFDSTGTL				AGAA
	250	260	270	280	290	300
	310	320	330	340	350	360
m097.pep	LGTSSTTPYVESAAGV					
os · · · pop					1:111111	1111
g097	LGTSSTTPYVESAAGV	SAGGRTGLTA	VTVGVLMLAC	LMFSPLAKSV	PVFATAPALL	YVGT
	310	320	330	340	350	360
	370	380	390	400	410	420
m097.pep	QMLRSARDIDWDDMTE	AAPAFLTIVE	MPFTYSIADG	IAFGFISYAV	VKLLCRRTKD	VPPM
g097	OMLRSARDIDWDDMTE					IIII
9097	370	380	390	400	410	420
	430					
m097.pep	VWIVAVLWALKFWYLG	X				
		1				
g097	VWVVAVLWALKFWYLG 430	iX.				
	430					

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 343> a097.seq

1	ATGGACACTT	CAAAACAAAC	ACTGTTGGAC	GGGATTTTTA	AGCTGAAGGC
51	AAACGGTACG	ACGGTGCGTA	CCGAGTTGAT	GGCGGGTTTG	ACAACTTTTT
101	TGACGATGTG	CTACATCGTT	ATCGTCAACC	CTCTGATTTT	GGGCGAGACC
151	GGCATGGATA	TGGGGGCGGT	ATTCGTCGCT	ACCTGTATCG	CGTCTGCCAT
201	CGGCTGTTTT	GTTATGGGTT	TTGTCGGCAA	CTATCCGATT	GCACTCGCAC
251	CGGGGATGGG	GCTGAATGCC	TATTTCACCT	TTGCCGTCGT	TAAGGGTATG
301	GGCGTGCCTT	GGCAGGTTGC	GTTGGGTGCG	GTGTTCATCT	CCGGTCTGAT
351	TTTCATCCTG	TTCAGCTTTT	TTAAAGTCAG	GGAAATGCTG	GTCAACGCAC
401	TGCCTATGGG	TTTGAAAATG	TCGATTGCTG	CCGGTATCGG	TTTGTTTTTG
451	GCACTGATTT	CCCTGAAAGG	CGCAGGCATT	ATCGTTGCCA	ATCCGGCAAC
501	CTTGGTCGGC	TTGGGCGATA	TTCATCAGCC	GTCCGCGTTG	TTGGCACTGT
551	TCGGTTTTGC	CATGGTGGTC	GTATTGGGAC	ATTTCCGCGT	TCAAGGCGCA
601	ATCATCATCA	CCATTTTGAC	GATTACCGTC	ATTGCCAGCC	TGATGGGTTT
651	GAACGAATTT	CACGGCATCA	TCGGCGAAGT	GCCGAGCATT	GCGCCGACTT
701	TTATGCAGAT	GGATTTTAAA	GGGTTGTTTA	CCGTCAGCAT	GGTCAGCGTG
751	ATTTTCGTCT	TTTTCCTAGT	CGATCTGTTC	GACAGTACCG	GAACACTGGT
801	CGGTGTATCG	CATCGTGCCG	GACTGCTGGT	GGACGGTAAG	CTGCCCCGCC
851	TGAAACGCGC	ACTGCTTGCA	GACTCTACCG	CTATTGTGGC	AGGTGCGGCT
901	TTGGGTACTT	CTTCAACCAC	GCCTTATGTG	GAAAGTGCGG	CGGGCGTATC
951	GGCAGGCGGG	CGGACAGGTC	TGACGGCGGT	TACCGTCGGC	GTATTGATGC
1001	TCGCCTGCCT	GATGTTTTCA	CCTTTGGCGA	AAAGTGTTCC	CGCTTTTGCC
1051	ACCGCGCCCG	CCCTGCTTTA	TGTCGGCACG	CAGATGCTCC	GCAGTGCGAG
1101	GGACATCGAT	TGGGACGATA	TGACGGAAGC	CGCACCCGCA	TTCCTGACCA
1151	TTGTCTTCAT	GCCGTTTACC	TATTCGATTG	CAGACGGCAT	CGCTTTCGGC
1201	TTCATCAGTT	ATGCCGTGGT	TAAACTTTTA	TGCCGCCGCA	CCAAAGACGT
1251	TCCGCCTATG	GTATGGATTG	TTGCCGTATT	GTGGGCACTG	AAATTCTGGT

BNSDOCID: <WO___9957280A2___>

1301 ATTTGGGCTG A

This corresponds to the amino acid sequence <SEQ ID 344; ORF 097.a>: a097.pep

1	MDTSKQTLLD	GIFKLKANGT	TVRTE <u>LMAGL</u>	TTFLTMCYTV	TUNDETT CET
51	GMDMG <u>AVFVA</u>	_TCIASAIGCF	VMGFVGNYPI	ALAPGMGLNA	YETEDWINCM
101	GVPWQVALGA	VFISGLIFIL	FSFFKVREML	VNAT.PMGT.KM	STANCICIET
151	ALISLKGAGI	IVANPATLVG	LGDIHQPSAL	LALEGEAMW	VIGHERVOCA
201	<u> </u>	_IASLMGLNEF	HGIIGEVPSI	APTEMOMDER	GI FTUSMUSU
251	IFVFFLVDLF	DSTGTLVGVS	HRAGLLVDGK	LPRIKRAT.T.A	DSTATUACAA
301	LGTSSTTPYV	ESAAGVSAGG	RTGLTAVTVG	VIMLACIMES	DIVECTOVEY
351	TAPALLYVGT	QMLRSARDID	WDDMTEAAPA	FITTVEMPET	VETADOTADO
401	FISYAVVKLL	CRRTKDVPPM	VWIVAVIWAT.	KFWYLC*	TOTADGIALG

m097/a097 99.3% identity in 436 aa overlap

430

	1.0					
m097.pep	MDTSKOTI I DOTEKI	20	30	40	50	60
mos, pep	MDTSKQTLLDGIFKI	KANGTTVR	TELMAGLTTFI	LTMCYIVIVN	PXILGETGMD	MGAVFVA
a097		Kancaara Kancaara	TELMACI MODE	T	1 11111	111111
	MDTSKQTLLDGIFKI	20	TELMAGLTIFI 30	TMCYIVIVN		
	20	20	30	40	50	60
	70	80	90	100	110	
m097.pep	TCIASAIGCFVMGFV		PGMGI.NAYETE	TOO	110	120
				11111111	4 3 4 1 4 4 1 1 1 4	
a097	TCIASAIGCFVMGFV	GNYPIALA	PGMGLNAYFTF	'AWWKGMGWD		
	70	80	90	100	WQVALGAVET:	
				200	110	120
	130	140	150	160	170	180
m097.pep	FSFFKVREMLVNALP	MGLKMSIA	AGIGLFLALIS	LKGAGIIVA	ים ביים עמו	
- 007	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1		1 1 1 1 1 1 1 1 1		
<b>a</b> 097	LOILINAKEMPANMPE	MOTUMOTAN	AGIGLELALIS	LKGAGIIVA	NPATLVGLGDI	HOPSAL
	130	140	150	160	170	180
	190	200				
m097.pep		200	210	220	230	240
mos/,pep	LALFGFAMVVVLGHF	KVQGAT111	TLTITVIASL	MGLNEFHGI	GEVPSIAPTF	MQMDFE
a097	I.AI.FGEAMWWW.CHE			1111111		11111:
	LALFGFAMVVVLGHFF	200	210	MGLNEFHGI]	GEVPSIAPTF	MQMDFK
	130	200	210	220	230	240
	250	260	270	280	200	
m097.pep	GLFTVSMVSVIFVFFI	VDLFDSTG	TIVGTSHRAGI	LUDCKI DDI	290	300
	* * *	1 1 1 1 1 1 1				
a097	GLFTVSMVSVIFVFFI	VDLFDSTG	TLVGVSHRAGI	LIVDGKLPRI	TITITITITI	
	250	260	270	280	290	300
					230	300
0.03	310	320	330	340	350	360
m097.pep	LGTSSTTPYVESAAGV	SAGGRTGL	TAVTVGVLMLA	CLMFSPLAK	SVPAFATAPA	
a097		1	1111111111	11111111		
a0 <i>91</i>	DG19911L1AE9WWGA	PAGGKIGE	TAVTVGVLMLA	CLMFSPLAK	SVPAFATAPA:	LLYVGT
	310	320	330	340	350	360
	370	380	200			
m097.pep			390	400	410	420
- 11-	QMLRSARDIDWDDMTE	UNENETIT.	VEMPETYSIAD	GIAFGFISY	AVVKLLCRRT	KDVPPM
a097	QMLRSARDIDWDDMTE	ΑΔΡΑΓΙΤΙ	/		<u> </u>	11111
	370	380	390	400		
			350	400	410	420
	430					
m097.pep	VWIVAVLWALKFWYLG:					
22-		l				
a097	VWIVAVLWALKFWYLG	<				
	430					

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 345>: g098.seq ATGACCGCCG ACGGTCTCTT CGTCGCTTTC AACTTCAATA CGTTTGCCGT 1 51 TGTGCGAATA TTGATACCAG TACAGCAGGA TGCTGCCCAG GCTGGCGATC 101 AGTTTGTCGG CGATGTCGCG CGCTTCGCTG TCGGGATGGC TTTCGCGTTC 151 GGGATGAACG CAGCCGAGCA TGGACACGCC GGTACGCATC ACGTCCATCG 201 GATGGGTATG TGCAGGCAGG CTTTCCAAAA CTTTAATCAC ACGGATAGGC 251 AGGCCGCGCA TGGATTTGAG CTTGGTTTTA TAAGCGGCCA GCTCGAATTT 301 GTTGGGCAGA TGGCCGTGAA TCAGCAAGTG TGCGACTTCT TCAAACTCGC 351 ATTTTTGTGC CAAATTAGAA TGTCGTAA This corresponds to the amino acid sequence <SEQ ID 346; ORF 098.ng>: g098.pep MTADGLFVAF NFNTFAVVRI LIPVQQDAAQ AGDQFVGDVA RFAVGMAFAF 51 GMNAAEHGHA GTHHVHRMGM CRQAFQNFNH TDRQAAHGFE LGFISGQLEF VGQMAVNQQV CDFFKLAFLC QIRMS* 101 The following partial DNA sequence was identified in N. meningitidis <SEQ ID 347>: m098.seq ATGACCGCCG ATGGTCTCTT CGTCGCTTTC AACCTCAATG CGTTTGCCGT 1 51 TGTGCGAATA TTGATACCAG TACAAGAGGA TGCTGCCGAG GCTGGCGATC 101 AGTTTGTCGG CGATGTCGCG CGCTTCACTT TCCGGATGGC TTTCACGTTC 151 AGGATGAACG CAGCCCAGCA TGGATACGCC GGTACGCATT ACGTCCATCG 201 GATGGGTATG TGCAGGCAGG CTTTCCAAAA CTTTAATCAC ACGGATAGGC 251 AGGCCGCGCA TGGATTTGAG CTTGGTTTTA TAAGCGGCCA GCTCGAATTT 301 GTTGGGCAGA TGGCCGTGAA TCAGCAGGTG GGCGACTTCT TCAAACTCGC 351 ATTTTGTGC CAAATCAGAA TGTCGTAA This corresponds to the amino acid sequence <SEQ ID 348; ORF 098>: m098.pep MTADGLFVAF NLNAFAVVRI LIPVQEDAAE AGDQFVGDVA RFTFRMAFTF 1 51 RMNAAQHGYA GTHYVHRMGM CRQAFQNFNH TDRQAAHGFE LGFISGOLEF VGQMAVNQQV GDFFKLAFLC QIRMS* Computer analysis of this amino acid sequence gave the following results: Homology with a predicted ORF from N. gonorrhoeae ORF 098 shows 89.6% identity over a 125 aa overlap with a predicted ORF (ORF 098.ng) from N. gonorrhoeae: m098/g098 10 20 30 40 m098.pep MTADGLFVAFNLNAFAVVRILIPVQEDAAEAGDQFVGDVARFTFRMAFTFRMNAAOHGYA MTADGLFVAFNFNTFAVVRILIPVQQDAAQAGDQFVGDVARFAVGMAFAFGMNAAEHGHA q098 10 20 30 40 50 60 90 100 110 GTHYVHRMGMCRQAFQNFNHTDRQAAHGFELGFISGQLEFVGQMAVNQQVGDFFKLAFLC m098.pep q098  ${\tt GTHHVHRMGMCRQAFQNFNHTDRQAAHGFELGFISGQLEFVGQMAVNQQVCDFFKLAFLC}$ 70 80 90 100 110 120 m098.pep QIRMSX 

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 349>: a098.seq

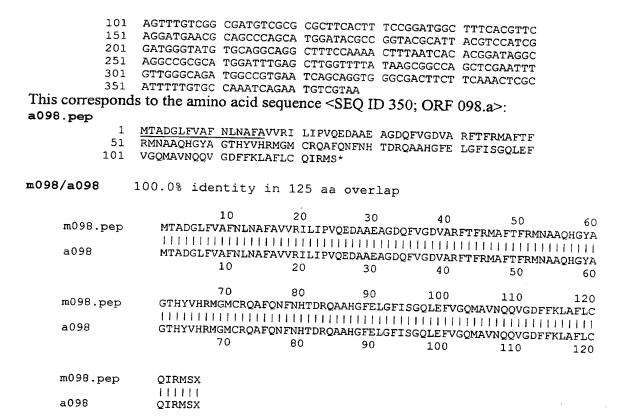
- 1 ATGACCGCCG ATGGTCTCTT CGTCGCTTTC AACCTCAATG CGTTTGCCGT
- 51 TGTGCGAATA TTGATACCAG TACAAGAGGA TGCTGCCGAG GCTGGCGATC

BNSDOCID: <WO__9957280A2_l_>

g098

QIRMSX

graphe grapher abla



# The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 351>:

1	ATGCTGGGAC GC	GCGTCCAT	GATGCGCCTG	CCCGATATTC	TOGGOGTOCA
51	GCTGACGGGC AA	ACGGCAGG	CGGGCATTAC	TGCCACAGAC	ATCGTGTTGG
101	CACTGACCGA AT	CTTGCGT	AAAGAGCGCG	TGGTCGGGG	CTTTCGIGIIGG
151	TTTTTCGGCG AGO	GCGCGAG	AAGCCTGTCT	ATCGGCGACC	GCCCCACCAM
201	TTCCAACATG ACC	CCGGAGT	TCGGCGCGAC	TGCCGCCATG	TTCCCCAT
251	ACGCGCAAAC TAT	TGATTAT	TTGAAACTGA	CCGGACGTGA	CCACCCCCA
301	GTGAAATTGG TGC	AAACCTA	CGCCAAAACC	GCAGGCTTAT	CCCCACGCGCAG
351	CTTGAAAACC GCC	CGTTTATC	CGCGCGTTTT	CAAATTTCAT	GGGCAGGTGG
401	TAACGCGCAA TAT	GGCAGGC	CCGAGCAACC	CGCACGCCCC	TIGAGCAGCG
451	GCCGATTTGG CGC	CGAAAGG	GCTGGCGAAG	CCTTACCAAC	TTTTGCCACC
501	CGGCCAAATG CCT	GACGGTG	CAGTGATTAT	TCCCCCCAAG	AGCCTTCAGA
551	CCAATACTTC CAA	CCCGCGC	AACGTTGTCC	CCCCCCCATT	ACTICGIGIA
601	AATGCCAACC GCC	TCGGCTT	GAAACGCAAA	CCGCCGCACT	GTTGGCACGC
651	TGCCCCGGGT TCA	AAAGTAG	CCGGAATCTA	CCIIGGGIGA	AATCTTCGTT
701	TGCCCGAAAT GGA	AAAACTC	CCCGGAATCIA	TITGAAAGAA	GCAGGCTTGT
751	ACCTGTAACG GCA	TGAGCGG	CGCGCTGGTA	CCCTCGCCTT	CGCATGTACC
801	CATCGACCGC GAt	ttataca	cGCGCTCGaC	CCGAAAATCC	AACAAGAAAT
851	TCGACGGCCG TAT	CCATCCC	TATCCCARA	ATTGTCAGGC	AACCGCAACT
901	CCTTTGGTCG TTG	CCTACCC	ATTICCGAAAC	AGGCTTTCCT	CGCTTCGCCT
951	CCTTTGGTCG TTG	CCIACGC	AT TGGCAGGT	AGCATCCGTT	TCGATATTGA
1001	AAACGACGTA CTC	CCAACAA	CAGACGGCCG	CGAAATCCGC	CTGAAAGATA
1051	TCTGGCCGAC AGA	CCCACAG	ATCGATGCCA	TCGTTGCCGA	ATATGTGAAA
1101		ACCCCCC	TTATATCCCG	ATGTCCGACA	CCGGCACAGC
1151	GCAAAAAGCA CCA	AGCCCCC	TGTACGACTG	GCGACCGATG	TCCACCTACA
1201	TCCGCCGTCC GCC	TATTGG	GAAGGCGCAC	TGGCAGGGGA	ACGTACATTA
1251	AGAGGTATGC GTC		GATTTTGCCC	GACAACATCA	CCACCGACCA
*577	CATCTCgcca tCC	AATGCGA	TTTTGGCCGG	cagTGCcgca	ggtgaATATT

***

```
1301 TGGCGAAAAT GGGTTTGCCT GAAGAagaCT TCAACTCTTA CGCAACCCAC
               CGCGGCGACC ACTTGACCGC CCAACGCGCA ACCTTCGCCA ATCCGAAACT
         1401 GTTTAACGAA ATGGTGAGAA ACGAAGACGG CAGCGTACGC CAAGGTtcgt
         1451 tggcacgcgT tgaacCAGAA GGCCAAACCA TGCGCATGTG GGAAGCCATC
         1501 GAAACCTATA TGAACCGCAA ACAGCCGCTT ATCATCATTG CCGGTGCGGA
         1551 CTATGGTCAA GGCTCAAGCC GCGACTGGGC GGCGAAGGGC GTGCGGCTGG
         1601 CGGGTGTGGA AGCCATCGCC GCCGAAGGTT TCGAGCGCAT CCACCGCACC
         1651 AACCTCATCG GCATGGGCGT CTTGCCGCTG CAATTCAAAC CCGGCACCAA
         1701 CCGCCATACC CTGCAACTGG ACGGTACGGA AACCTACGAC GTTGTCGGCG
         1751 AACGCACACC GCGCTGCGGC CTGACCCTCG TGATTCACCG TAAAAACGGA
         1801 GAAACCGTCG AAGTTCCGGT TACCTGCCGC CCCGATACCG CAGAAGAAGC
         1851 ATTGGTATAT GAAGCCGGCG GCGTATTGCA ACGGTTTGCA CAGGACTTTT
         1901 TGGAAGGGAA CGCGGCTTAG
This corresponds to the amino acid sequence <SEQ ID 352; ORF 099.ng>:
     g099.pep
              MLGRASMMRL PDIVGVELTG KROAGITATD IVLALTEFLR KERVVGAFVE
            1
              FFGEGARSLS IGDRATISNM TPEFGATAAM FAIDAOTIDY LKLTGRDDAO
           51
          101 VKLVETYAKT AGLWAGGLKT AVYPRVLKFD LSSVTRNMAG PSNPHARFAT
          151 ADLAAKGLAK PYEEPSDGQM PDGAVIIAAI TSCTNTSNPR NVVAAALLAR
          201 NANRLGLKRK PWVKSSFAPG SKVAGIYLKE AGLLPEMEKL GFGIVAFACT
              TCNGMSGALD PKIQQEIIDR DLYATAVLSG NRNFDGRIHP YAKQAFLASP
               PLVVAYALAG SIRFDIENDV LGVADGREIR LKDIWPTDEE IDAIVAEYVK
          351 PQQFRDIYIP MSDTGTAQKA PSPLYDWRPM STYIRRPPYW EGALAGERTL
          401 RGMRPPAILP DNITTDHISP SNAILAGSAA GEYLAKMGLP EEDFNSYATH
               RGDHLTAQRA TFANPKLFNE MVRNEDGSVR QGSLARVEPE GQTMRMWEAI
               ETYMNRKOPL IIIAGADYGQ GSSRDWAAKG VRLAGVEAIA AEGFERIHRT
          551 NLIGMGVLPL QFKPGTNRHT LQLDGTETYD VVGERTPRCG LTLVIHRKNG
          601 ETVEVPVTCR PDTAEEALVY EAGGVLQRFA QDFLEGNAA*
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 353>:
     m099.seq
               ATGCTGGGAC GCGCGTCCAT GATGCGCCTG CCCGATATTG TCGGCGTTGA
               GCTGAACGGC AAACGGCAGG CGGGCATTAC GGCGACGGAT ATTGTGTTGG
               CACTGACCGA GTTTCTGCGC AAAGAACGCG TGGTCGGGGC GTTTGTCGAA
               TTCTTCGGCG AGGGCGCGAG AAGCCTGTCT ATCGGCGACC GCGCGACCAT
               TTCCAACATG ACGCCGGAGT TCGGCGCGAC TGCCGCGATG TTCGCTATTG
          201
          251 ATGAGCAAAC CATTGATTAT TTGAAACTGA CCGGACGCGA CGACGCGCAG
          301 GTGAAATTGG TGGAAACCTA CGCCAAAACC GCAGGCTTGT GGGCAGATGC
          351 CTTGAAAACC GCCGTTTATC CTCGCGTTTT GAAATTTGAT TTGAGCAGCG
          401 TAACGCGCAA TATGGCAGGC CCAAGTAACC CGCATGCCCG TTTTGCGACC
          451 GCCGATTTGG CGGCGAAAGG GCTGGCGAAG CCTTACGAAG AGCCTTCGGA
          501 CGGCCAAATG CCCGACGGCT CGGTCATCAT CGCCGCGATT ACCAGTTGCA
          551 CCAACACTTC CAACCCGCGC AACGTTGTTG CCGCCGCGCT CTTGGCACGC
          601 AATGCCAACC GTCTCGGCTT GAAACGCAAA CCTTGGGTGA AATCTTCGTT
               TGCCCCGGGT TCAAAAGTAG CCGAAATCTA TTTGAAAGAA GCGGGCCTGT
          651
               TGCCCGAAAT GGAAAAACTC GGCTTCGGTA TCGTCGCCTT CGCCTGCACC
               ACCTGCAACG GCATGAGTGG CGCGCTGGAT CCGAAAATCC AGAAAGAAAT
               CATCGACCGC GATTTGTACG CCACCGCCGT ATTATCAGGC AACCGCAACT
          851 TCGACGCCG TATCCACCCG TATGCGAAAC AGGCTTTCCT CGCTTCGCCT
          901 CCGTTGGTCG TTGCCTACGC GCTGGCAGGC AGTATCCGTT TCGATATTGA
          951 AAACGACGTA CTCGGCGTTG CAGACGGCAA GGAAATCCGC CTGAAAGACA
         1001 TTTGGCCTGC CGATGAAGAA ATCGATGCCG TCGTTGCCGA ATATGTGAAA
         1051 CCGCAGCAGT TCCGCGATGT GTATGTACCG ATGTTCGACA CCGGCACAGC
         1101 GCAAAAGCA CCCAGTCCGC TGTACGATTG GCGTCCGATG TCCACCTACA
         1151 TCCGCCGTCC GCCTTACTGG GAAGGCGCGC TGGCAGGGGA ACGCACATTA
         1201 AGAGGTATGC GTCCGCTGGC GATTTTGCCC GACAACATCA CCACCGACCA
         1251 CCTCTCGCCG TCCAATGCGA TTTTGGCCGT CAGTGCCGCA GGCGAGTATT
         1301 TGGCGAAAAT GGGTTTGCCT GAAGAAGACT TCAACTCTTA CGCAACCCAC
         1351 CGCGGCGACC ACTTGACCGC CCAACGCGCT ACCTTCGCCA ATCCGAAACT
```

1401 GTTTAACGAA ATGGTGAAAA ACGAAGACGG CAGCGTGCGC CAAGGCTCGT 1451 TCGCCCGCGT CGAACCCGAA GGCGAAACCA TGCGCATGTG GGAAGCCATC 1501 GAAACCTATA TGAACCGCAA ACAGCCGCTC ATCATCATTG CCGGTGCGGA

BNSDOCID: <WO___9957280A2_I_>



This corresponds to the amino acid sequence <SEQ ID 354; ORF 099>:

```
1 MLGRASMMRL PDIVGVELNG KRQAGITATD IVLALTEFLR KERVVGAFVE
51 FFGEGARSLS IGDRATISNM TPEFGATAAM FAIDEQTIDY LKLTGRDDAQ
101 VKLVETYAKT AGLWADALKT AVYPRVLKFD LSSVTRNMAG PSNPHARFAT
151 ADLAAKGLAK PYEEPSDGQM PDGSVIIAAI TSCTNTSNPR NVVAAALLAR
201 NANRLGLKRK PWVKSSFAPG SKVAEIYLKE AGLLPEMEKL GFGIVAFACT
251 TCNGMSGALD PKIQKEIIDR DLYATAVLSG NRNFDGRIHP YAKQAFLASP
301 PLVVAYALAG SIRFDIENDV LGVADGKEIR LKDIWPADEE IDAVVAEYVK
351 PQQFRDVYVP MFDTGTAQKA PSPLYDWRPM STYIRRPPYW EGALAGERTL
401 RGMRPLAILP DNITTDHLSP SNAILAVSAA GEYLAKMGLP EEDFNSYATH
451 RGDHLTAQRA TFANPKLFNE MVKNEDGSVR QGSFARVEPE GETMRMWEAI
501 ETYMNRKQPL IIIAGADYGQ GSSRDWAAKG VRLAGVEAIV AEGFERIHRT
551 NLIGMGVLPL QFKPDTNRHT LQLDGTETYD VVGERTPRCD LTLVIHRKNG
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae

ORF 099 shows 96.2% identity over a 639 aa overlap with a predicted ORF (ORF 099.ng) from N. gonorrhoeae:

m099/g099

	10	20	30	40	50	<b>CO</b>
m099.pep	MLGRASMMRLPDIV	GVELNGKRO	AGITATDIVL	ALTEFLEKED	Marresea Marresea	60
		1111:11	1 1 1 1 1 1 1 1 1	111111111		1111111
g099	MLGRASMMRLPDIV	GVELTGKRO	AGTTATITUL:	יים מאר האת המארות אני מאר האת המארות		111111
	10	20	30	HUIEFLRKER		EGARSLS
		2,0	30	40	50	60
	70	80				
m099.pep			90	100	110	120
	IGDRATISNMTPEF	GATAAMFAI	DEQTIDYLKL	TGRDDAQVKL1	/ETYAKTAGL	WADALKT
9099		1   1   1   1   1   1	1 1 4 4 1 1 1 1 1			
9023	IGDRATISNMTPEF	GATAAMFAI:	DAQTIDYLKLI	GRDDAOVKL	ETYAKTAGIJ	MAGGI 200
	70	80	90	100	110	120
					110	120
	130	140	150	160	170	
m099.pep	AVYPRVLKFDLSSV	TRNMAGPSN	א זרו גידי מים PHAPE איי	A VOL A VOVO	170	180
				MUGUAKPYEE	PSDGOMPDGS	IAAIIV
g099	AVYPRVIKEDISSV	IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII		1111111111	11111111:	
_	AVYPRVLKFDLSSV	140	PHARFATADLA	AKGLAKPYEE	PSDGQMPDGA	VIIAAI
	130	140	150	160	170	180
	100					
m099.pep	190	200	210	220	230	240
moss.pep	TSCTNTSNPRNVVA	ALLARNANI	RLGLKRKPWVK	SSFAPGSKVA	EIYLKEAGII	DEMENT
- 0.0-				11111111		11171
g099	TSCTNTSNPRNVVA	ALLARNANE	LGLKRKPWVK	SSFAPGSKVA	 	71111
	190	200	210	220		
				220	230	240
	250	260	270			
m099.pep		בעת המצו מאות המצו	27U	280	290	300
	GFGIVAFACTTCNGM		VETIDEDLAY	PAVLSGNRNF)	OGRIHPYAKQ.	AFLASP
g099		11111111	:	1   1   1   2   1   4		
3.00	GFGIVAFACTTCNGM	SGALDPKIQ	QEIIDRDLYA	TAVLSGNRNFI	GRIHPYAKO	AFTASP
	250	260	270	280	290	300
						300
	310	320	330	340	350	2.50
				- 10	350	360

m099.pep	PLVVAYALAG	SIRFDIENDV	LGVADGKEIRI	LKDIWPADEE	DAVVAEYVKE	QQFRDVYVP
						1         :   :
g099	PLVVAYALAG 310		LGVADGREIRI 330	LKDIWPTDEE: 340		
	310	320	330	340	350	360
	370	380	390	400	410	420
m099.pep			STYIRRPPYW	EGALAGERTLF	RGMRPLAILPD	NITTDHLSP
g099	_		STYIRRPPYW			
	370	380	390	400	410	420
	430	4.40	450	460	470	480
m099.pep			EEDFNSYATH			
oss.pep						:
g099	SNAILAGSAA	GEYLAKMGLP	EEDFNSYATH	RGDHLTAORAT	FANPKLFNEM	
	430	440	450	460	470	480
	490	500		520	530	540
m099.pep			ETYMNRKQPLI	IIIAGADYGQG	SSRDWAAKGV	
g099	OGSTARVERE		ETYMNRKQPLI		ון ון ון ון ון ון	:
9000	490	500		520	530 530	540
						3.10
	550	560	570	580	590	600
m099.pep	AEGFERIHRT	NLIGMGVLPL	QFKPDTNRHTI	LOLDGTETYDV	VGERTPRCDL	TLVIHRKNG
g099	AEGFERIHRT 550	NLIGMGVLPL 560	QFKPGTNRHTI 570			
	550	560	570	580	590	, 600
	610	620	630	640		
m099.pep			EAGGVLORFAC			
		:	1111111111	11111111		
g099			<b>EAGGVLQRFA</b> C	DFLEGNAAX		
	610	620	630	640		

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 355>: a099.seq

1	ATGCTGGGAC	GCGCGTCCAT	GATGCGCCTG	CCCGATATTG	TCGGCGTTGA
51	GCTGAACGGC	AAACGGAAGG	CGGGCATTAC	GGCGACGGAT	ATTGTGTTGG
101	CACTGACCGA	GTTTCTGCGC	AAAGAACGCG	TGGTCGGGGC	GTTTGTCGAA
151	TTCTTCGGCG	AGGGCGCGAG	AAGCCTGTCT	ATCGGCGACC	GCGCGACCAT
201	TTCCAACATG	ACGCCGGAGT	TCGGCGCGAC	TGCCGCGATG	TTCGCTATTG
251	ATGAGCAAAC	CATTGATTAT	TTGAAACTGA	CCGGACGCGA	CGACGCGCAG
301	GTGAAATTGG	TGGAAACCTA	CGCCAAAACC	GCAGGCTTGT	GGGCAGATGC
351	CTTGAAAACC	GCCGTTTATC	CGCGCGTTTT	GAAATTTGAT	TTGAGCAGCG
401	TAACGCGCAA	TATGGCAGGC	CCGAGCAACC	CGCACGCGCG	TTTTGCGACC
451	GCCGATTTGG	CCGGCAAAGG	CTTGGCTAAA	CCTTACGAAG	AGCCTTCAGA
501	CGGCCAAATG	CCTGACGGTG	CAGTGATTAT	TGCCGCGATT	ACTTCCTGTA
551	CCAATACTTC	CAATCCGCGC	AACGTTGTCG	CCGCCGCGCT	GTTGGCACGC
601	AATGCCAACC	GCCTCGGCTT	GCAACGCAAA	CCTTGGGTGA	AATCTTCGTT
651	TGCCCCGGGT	TCAAAAGTAG	CCGAAATCTA	TTTGAAAGAA	GCAGATCTGC
701	TGCCCGAAAT	GGAAAAACTC	GGCTTCGGTA	TCGTTGCCTT	CGCATGTACC
751	ACCTGTAACG	GCATGAGCGG	CGCGCTGGAT	CCGAAAATCC	AGAAAGAAAT
801	CATCGACCGC	GATTTGTACG	CCACCGCCGT	ATTGTCAGGC	AACCGCAACT
851	TTGACGGCCG	TATCCATCCG	TATGCGAAAC	AGGCTTTCCT	CGCTTCGCCT
901	CCGTTGGTCG	TTGCCTACGC	GCTGGCAGGC	AGCATCCGTT	TCGATATTGA
951	AAACGACGTA	CTCGGCGTTG	CAGACGGCAA	AGAAATCCGC	CTGAAAGACA
1001	TTTGGCCTAC	CGATGAAGAA	ATCGATGCCA	TCGTTGCCGA	ATATGTGAAA
1051	CCGCAGCAAT	TTCGCGACGT	TTATATCCCG	ATGTTCGACA	CCGGCACAGC
1101	GCAAAAAGCA	CCAAGCCCGC	TGTACGACTG	GCGTCCAATG	TCTACCTATA
1151	TCCGCCGCCC	ACCTTACTGG	GAAGGCGCAC	TGGCAGGGGA	ACGCACATTA
1201	AGCGGTATGC	GTCCGCTGGC	GATTTTGCCC	GACAACATCA	CCACCGACCA

1251	TCTCTCGCCA TCCAATCCCA TTTTTCCCAACCAAC
1301	TITIOGCAAG CAGTGCTATA
1351	CGTGGCGACC ACTTGACCGC CCAACGCGCA ACCTTCGCCA ATCCGAAACT
1401	GTTTAACGAA ATGGTGAGAA ACGAAGACGG CAGCGTACGC CAAGGTTCGC
1451	TGGCACGCGT TGAACCCGAA GGCCAAACCA TGCGCATGTG GGAAGCCATC
1501	GAAACCTATA TGAACCGCAA ACAGCCGCTC ATCATCATTG CCGGCGCGGA
1551	CIACGGTCAA GGCTCAAGCC GCGACTGGGC TGCAAAAGCC CTACCCCTCC
1601	CCGCGTGGA AGCGATTGTT GCCGAAGGCT TCGAGCGTAT CCAGCGGAGG
1651	AACIIGAICG GTATGGGCGT GTTGCCGCTG CACTTCAAAC CCCCTT CCTT
1701	CCGCCACACC CTGCAACTGG ACGGTACGGA AACCTACGAC CTTCTCCCC
1751	AACGCACACC GCGCTGCGAC CTGACCCTTG TCATTCACCC TAAAAAACCCC
1801	GAGACCGICG AAGTCCCCAT TACCTGCCGC CTCCATACCC CACAACAACA
1851	GIIGGIAIAI GAAGCCGGTG GCGTATTGCA ACGGTTTCCA CACGAMMUMM
1901	I GGAAGGAA CGCGGCTTAG
inis correspond	ds to the amino acid sequence <seq 099.a="" 356;="" id="" orf="">:</seq>
a099. <b>pe</b> p	
1	MLGRASMMRL PDIVGVELNG KRKAGITATD IVLALTEFLR KERVVGAFVE
51	FIGEGARSLS IGDRATISNM TPEFGATAAM FAIDFOTIDY IVINCEDED.
101	VKDVETIAKI AGLWADALKI AVYPRVIKED I SSVTDNMAC DENDUADERE
151	ADDAGRGLAR PILEPSDGOM PDGAVITAAT TSCTNTSNDD AUGUARRET
201	NANKLGLOKK PWVKSSEAPG SKVAETYLKE ADIIDEMENT CECTURER OF
251	TONGMOGRAD PRIORELIDE DLYATAVISC NEWFOCETUR VAKCARIAGE
301	IDVVAIADAG SIREDIENDV LGVADGKETR IKDIMPTDER IDDITATION
351	FORFAUVILE MEDICIAOKA PSPLYDWRPM STYTEREDDVW FOALAGERT
401	SGMRPLAILP DNITTDHLSP SNATLASSAA GEVLARMOLD EEDDWARD
451 501	AGUNTIAURA TEANPELENE MURNEDGSUR OCGIARUERE COMMENTERS
551	DITIMARY I TITAGADIGO GSSROWAAKC VPIACUEATU ABGEED TITA
601	NDIGHGVEFE QERPGINKHT LOLDGTETYD WWGFDTDDCD INLUTUDIO
001	ETVEVPITCR LDTAEEVLVY EAGGVLQRFA QDFLEGNAA*
-000/-000	08. ***
m099/a099	97.5% identity in 639 aa overlap
	10 20 30 40 50 60
m099.pep	MLGRASMMRLPDIVGVELNGKROAGITATDIVLALTEFIPERINGA TURBURA
	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
m099.pep a099	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
	MLGRASMMRLPDIVGVELNGKROAGITATDIVLALTEFIPERINGA TURBURGA
	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
a099	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
a099	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
a099 m099.pep	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
a099 m099.pep	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
a099 m099.pep a099	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
a099 m099.pep	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
a099 m099.pep a099 m099.pep	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
a099 m099.pep a099	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
a099 m099.pep a099 m099.pep	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
a099 m099.pep a099 m099.pep	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
a099 m099.pep a099 m099.pep a099	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
a099 m099.pep a099 m099.pep	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
m099.pep a099 m099.pep a099	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
a099 m099.pep a099 m099.pep a099	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
m099.pep a099 m099.pep a099	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
m099.pep a099 m099.pep a099	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
m099.pep a099 m099.pep a099	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
a099 m099.pep a099 m099.pep a099	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
a099 m099.pep a099 m099.pep a099	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
a099 m099.pep a099 m099.pep a099 m099.pep a099	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
a099 m099.pep a099 m099.pep a099 m099.pep a099	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
m099.pep a099 m099.pep a099 m099.pep a099 m099.pep a099	MLGRASMMRLPDIVGVELNCKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS
a099 m099.pep a099 m099.pep a099 m099.pep a099	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS

in electric electric

a099					 EYVKPQQFRD 350	
	370	380	390	400	410	420
m099.pep	MFDTGTAQKAPSPLYD	WRPMSTYIRR	PPYWEGALAG	ERTLRGMRPL	AILPDNITTD:	HLSP
a099	MFDTGTAQKAPSPLYD					HLSP
	370	380	390	400	410	420
	430	440	450	460	470	
m099.pep	SNAILAVSAAGEYLAK				470	480
moss.pep		11111111			:	
a099	SNAILASSAAGEYLAK					
4033	430	440	450	460	470	480
						100
	490	500	510	520	530	540
m099.pep	QGSFARVEPEGETMRM	WEAIETYMNR	KQPLIIIAGA:	DYGQGSSRDW.	AAKGVRLAGVI	EAIV
	-       :					
a099	QGSLARVEPEGQTMRM				<b>AAKG</b> VRLAGVI	EAIV
	490	500	510	520	530	540
	550	560	570	E00	F.00	<b>600</b>
m099.pep	AEGFERIHRTNLIGMG			580 ETYDINICEDT	590	600
moss.pep						
a099	AEGFERIHRTNLIGMG					
	550	560	570	580	590	600
	610	620	630	640		
m099.pep	ETVEVPVTCCLDTAEE					
a099	ETVEVPITCRLDTAEE					
	610	620	630	640		

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 357>: g102.seq

1 Atatacacaca actatacaca acacacata atatacacaga

1 Atatacacaca actatacaca acacacata atatacacaga

1			gctcttcggc		
51	gaaggttatc	ggcgcAGgta	tgttccccaa	ccccaccgcc	aacttggggg
101	acgggttaat	aggctcgctg	attgtgctgc	tgtacacctg	gtttccattc
151			tttggaagtc		
201	ggcaAGtttt	gacaccATGg	tcAAagacct	gctcgGaCGc	ggctggaaca
251	tcatcaacgg	catcgccgtc	gctttggTCc	tatacggctc	gacctacgcg
301	tacattttag	tcggcggtga	cctGACCGCC	AAAGGCAtcg	GCAgCGCAGT
351	AGGCGGCAAA	ATTTCgctca	CCGTCGGACA	actcgtcttc	<b>tTCGGCATCC</b>
401	TCGCCTTTTG	CGTATGGGCA	TCCGCACGCT	TGGTCGACCG	CTTTACCGGC
451	GTCCTCATCG	GCGGCATGGT	ATTAACCTTT	ATTTGGGCAA	CCGGCGGCCT
501	GGTTGCCGAT	GCCAAACCGT	CCGTCCTCTT	CGACACCCAA	GCCCCCGTCG
551	GCACCGGCTA	CTGGATTTAC	GCCGCCACCG	CCCTGCCCGT	CTGCCTCGCT
601	TCCTTCGGCT	TCCACGGCAA	CGTTTCCAGC	CTGCTCAAAT	ACTTTAAAGG
651	CGACGcgcCc	aaagtGgCGA	aATCcatctg	gGcaggtaca	ttggTTGCCt
701	tggtaattta	cgtccTCTgg	caaaccgcca	tCcaaagcaa	ccTGCcgcgc
751	aacgagttcg	cCCCcgtgat	tgccgccgag	aggcaactCT	CCGTCCTgaa
801	tgaaacccTG	tccaaattcg	cccaaaccgg	cgatatggat	aAaatattgt
851	ccctatttcc	ctacatggca	atcgccacct	cctttttagg	cgTAACctta
901	ggcctgtttg	acaacatcgc	cgacatcttc	aaatggaacg	acagtatgtc
951	cgggcggggc	accaaaaccg	tcgcgctgaa	cttcctgccg	CCCCtgattt
1001	cctggctgct	cctccccacc	ggcttcttta	ccgccattgg	tgcgtccggc
1051	ctggcggcaa	ccgtctggga	ccaagGcatc	atccccgcca	tgctgctcta
1101			gcGcaggcaa		
1151	gcttgtggct	gatgttagtc	ttccttttcg	gcatcgccaa	catcgccgca
1201	CAGGTATTGA	GccaAatgGa	ACtcgtCccc	GTATTTAAAG	GATAA

This corresponds to the amino acid sequence <SEQ ID 358; ORF 102.ng>: g102.pep

BNSDOCID: <WO___9957280A2_J_>

```
1 MSAKTPSLFG GAMIIAGKVI GAGMFPNPTA NLGDGLIGSL IVLLYTWFPF
51 SSGALMILEV NTHNPRGASF DTMVKDLLGR GWNIINGIAV ALVLYGSTYA
101 YILVGGDLTA KGIGSAVGGK ISLTVGQLVF FGILAFCVWA SARLVDRFTG
151 VLIGGMVLTF IWATGGLVAD AKPSVLFDTQ APVGTGYWIY AATALPVCLA
201 SFGFHGNVSS LLKYFKGDAP KVARSIWAGT LVALVIYVLW QTAIQSNLPR
251 NEFAPVIAAE RQLSVLNETL SKFAQTGDMD KILSLFPYMA IATSFLGVTL
301 GLFDNIADIF KWNDSMSGRG TKTVALNFLP PLISWLLLPT GFFTAIGASG
351 LAATVWDQGI IPAMLLYVSP QKIGAGKTYK VYGGLWLMLV FLFGIANIAA
401 QVLSQMELVP VFKG*
```

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 359>: m102.seq

```
ATGCCCAACA AAACCCCTTC ACTGTTCGGC GGCGCGATGA TTATCGCCGG
      CACGGTCATC GGCGCAGGCA TGCTCGCCAA CCCGACCGCC ACATCCGGCG
      TATGGTTTAC CGGCTCGCTG GCCGTGTTGC TGTACACCTG GTTTTCTATG
  101
 151 CTTTCCAGCG GCCTGATGAT TTTGGAAGTC AACACCCATT ATCCGCACGG
201 CGCAAGTTTC GACACGATGG TCAAAGACCT GCTCGGACGC GGCTGGAACA
      TCATCAACGG CATCGCCGTC GCCTTCGTTT TATACCTGCT TACTTACGCT
 301 TATATCTTCG TCGGCGGCGA CCTGACCGCC AAAGGCTTAG GCAGCGCGGC
 351 AGGCGGCGAC GTTTCACTCA CCGTCGGACA ACTCGTCTTC TTCGGCATCC
 401 TCGCCTTTTG CGTATGGGCA TCCGCACGCT TGGTCGACCG CTTCACCGGC
 451 GTCCTTATCG GCGGCATGGT ATTGACCTTT ATTTGGGCGG CCGGCGGGCT
      GATTGCCGAT GCCAAGCCGT CCGTCCTCTT CGATACCCAA GCCCCCGCCG
 551 GCACAAACTA CTGGATTTAC GCCGCCACCG CCCTGCCCGT CTGCCTCGCT
 601
      TCCTTCGGCT TCCACGGCAA CGTCTCCAGC CTGCTCAAAT ACTTTAAAGG
      CGACGCGCCC AAAGTGGCTA AATCCATCTG GACGGGCACA CTGATTGCGC
 651
      TGGTAATTTA CGTCCTCTGG CAAACCGCCA TCCAAGGCAA CCTGCCGCGC
 701
 751
      AACGAGTTCG CCCCCGTCAT CGCCGCCGAA GGGCAAGTCT CCGTCCTCAT
 801 CGAAACCCTG TCCAAATTCG CCCAAACCGG CAATATGGAC AAAATATTGT
 851 CCCTGTTTTC CTATATGGCG ATCGCCACCT CGTTTTTAGG CGTAACGCTC
 901 GGACTCTTCG ACTACATCGC CGACATCTTC AAATGGAACG ACAGCATCTC
     CGGCCGCACC AAAACCGCCG CGCTGACCTT CCTGCCGCCC CTGATTTCCT
 951
1001
      GCCTGCTCTT CCCCACCGGC TTCGTTACCG CCATCGGCTA CGTCGGCCTG
1051 GCGGCAACCG TCTGGACAGG CATCATCCCC GCCATGCTGC TCTACCGTTC
1101
     GCGCAAAAA TTCGGCGCAG GCAAAACCTA TAAAGTTTAC GGCGGCTTGT
     GGCTGATGGT TTGGGTCTTC CTTTTCGGCA TCGTCAACAT CGCCGCACAG
1151
1201 GTATTGAGCC AAATGGAACT CGTCCCCGTA TTTAAAGGAT AA
```

# This corresponds to the amino acid sequence <SEQ ID 360; ORF 102>: m102.pep..

```
1 MPNKTPSLFG GAMIIAGTVI GAGMLANPTA TSGVWFTGSL AVLLYTWFSM
51 LSSGLMILEV NTHYPHGASF DTMVKDLLGR GWNIINGIAV AFVLYLLTYA
101 YIFVGGDLTA KGLGSAAGGD VSLTVGQLVF FGILAFCVWA SARLVDRFTG
151 VLIGGMVLTF IWAAGGLIAD AKPSVLFDTQ APAGTNYWIY AATALPVCLA
201 SFGFHGNVSS LLKYFKGDAP KVAKSIWTGT LIALVIYVLW QTAIQGNLPR
251 NEFAPVIAAE GQVSVLIETL SKFAQTGNMD KILSLFSYMA IATSFLGVTL
301 GLFDYIADIF KWNDSISGRT KTAALTFLPP LISCLLFPTG FVTAIGYVGL
351 AATVWTGIIP AMLLYRSRKK FGAGKTYKVY GGLWLMVWVF LFGIVNIAAQ
401 VLSQMELVPV FKG*
```

m102/g102 86.0% identity in 415 aa overlap

m102.pep	10 MPNKTPSLFGGAM	20 IIAGTVIGAGI	30 MLANPTATSGV	40 WFTGSLAVI	50 LYTWFSMLSS	60 GLMILEV
g102	MSAKTPSLFGGAM		):     :	: 111 11	11111	
	10	20	30	40	50	60
m102.pep	70	80	90	100	110	120
mioz.pep	NTHYPHGASFDTM	KDLLGRGWNI	INGIAVAFVL	LLTYAYIF	VGGDLTAKGLG	SAAGGD
~1.00	_		11111111111	11111.	I I I I I I I I I I I I I I I I I I I	11.11
g102	NTHNPRGASFDTM	KDLLGRGWNI	INGIAVALVL	GSTYAYIL	VGGDLTAKGIG	SAVGGK
	70	80	90	100	110	120
m102.pep	130	140	150	160	170	180
roz.pep	VSLTVGQLVFFGII	AFCVWASARL	VDRFTGVLIGG	MVLTFIWA	AGGLIADAKPS	VLFDTO
g102	:	11111111		111111	. 1 1 1 . 1 ( ) ( )	1 1 1 1 1 1
	130	140	150	160		
			130	100	170	180
m102.pep	190	200	210	220	230	240
	APAGTNYWIYAATA	LPVCLASFGF	HGNVSSLLKYF	KGDAPKVAI	KSIWTGTLIAL	VIYVLW

g102	:  :     APVGTGYWIY 190			1	:   :    SIWAGTLVAL   230	 VIYVLW 240
m102.pep	11111:1111	NEFAPVIAAEGQV	ін анні	11:11111	1 1111111	HHH
g102	QTATQSNLPR 250	NEFAPVIAAERQI 260	270	280 280	290	300
100	310		330	340	350	than cr
m102.pep	GTEDATADIE	KWNDSISGR-TKT 	:  :		IIIII IIII	11
g102		KWNDSMSGRGTKT			IGASGLAAT	
	310	320	330	340	350	360
	360 3	70 380	. 390	400	410	
m102.pep	IPAMLLYRSR	KKFGAGKTYKVYG	GLWLMVWVFLFG	SIVNIAAQVLS	QMELVPVFK	GX
g102	IPAMLLYVSP 370	QKIGAGKTYKVYG 380			QMELVPVFK 410	GX

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 361>: a102.seq

```
1 ATGCCCACCA AAACCCCTTC ACTGTTCGGC GGCGCGATGA TTATCGCCGG
     CACGNTCATC GGCGCAGGTA TGCTCGCCAA CCCGACCGCC ACATCCGGCG
  51
      TATGGTTTAC CGGCTCGCTG GCCGTGTTGC TGTACACCTG GTTTTCCATG
 101
     CTCTCCAGCG GCCTGATGAT TTTGGAAGTC AACACCCACT ACCCCCACGG
 151
     CGCGANCTTC GACACCATGG TTAAAGACCT GCTCGGACGG AGCTGGAACA
 251 TCATCAACGG CATCGCCGTC GCCTTCGTTT TATACCTGCT TACTTACGCT
 301 TATATCTTCG TCGGCGGCGA CCTGACCGCC AAAGGCTTAG GCAGCGCGGC
      AGGCGGCAAT GTTTCACTCA CCGTCGGACA ACTCGTCTTC TTCGGCATTC
 401 TCGCCTTTTG CGTATGGGCA TCCGCACGCT TGGTCGACCG ATTCACCAGC
 451 GTCCTCATCG GCGGCATGGT ATTAACCTTT ATTTGGGCAA CCGGCGGCCT
 501 GATTGCCGAT GCCAAACTGC CCGTCCTCTT CGACACCCAA GCCCCTACCG
 551 GCACCAACTA CTGGATTTAT GTCGCCACCG CCCTGCCCGT CTGCCTTGCG
     TCATTCGGTT TCCACGGCAA CGTCTCCAGC CTGCTCAAAT ACTTTAAAGG
 651
     CGACGCGCC AAAGTGGCTA AATCCATCTG GACGGGCACA CTGATTGCGC
 701 TGGTAATTTA CGTCCTCTGG CAAACCGCCA TCCAANGCAA CCTGCCGCGC
 751 AACGAGTTCG CCCCGTGAT TGCCGCCGAA GGGCAAGTCT CCGTCNTGAT
 801 TGAAACCCTG TCCAAATTCG CCCAAACCGG CAATATGGAC AAAATATTGT
851 CCCTGTTTTC CTATATGGCG ATCGCCACCT CGTTTTTAGG CGTAACGCTC
901 GGACTCTTCG ACTACATCGC CGACATCTTC AAATGGAACG ACAGCGTGTC
 951 CGGCCGCACC AAAACCGCCG CGCTGACCTT CCTGCCGCCT NTAATTTCCT
1001 GCCTGCTCTT CCCCACCGC TTTGTTACCG CCATCGGNTA CGTCGGCCTG
1051 GCGGCAACCG TCTGGACAGG CATCATCCCC GCCATGCTGC TNTACCGTTC
1101
     GCGCAAAAA TTCGGCGCAG GCAAAACCTA TAAAGTTTAC GGCGGCTTGT
     GGCTGATGGT TTGGGTCTTC CTTTTCGGCA TCNTCAACAT CGCCGCACAN
1151
     GTATTGAGCC AAATGGAACT CGTCCCCGTA TTTAAAGGAT AA
1201
1202
```

This corresponds to the amino acid sequence <SEQ ID 362; ORF 102.a>: a102.pep

```
1 MPTKTPSLFG GAMIIAGTXI GAGMLANPTA TSGVWFTGSL AVLLYTWFSM
51 LSSGLMILEV NTHYPHGAXF DTMVKDLLGR SWNIINGIAV AFVLYLLTYA
101 YIFVGGDLTA KGLGSAAGGN VSLTVGQLVF FGILAFCVWA SARLVDRFTS
151 VLIGGMVLTF IWATGGLIAD AKLPVLFDTQ APTGTNYWIY VATALPVCLA
201 SFGFHGNVSS LLKYFKGDAP KVAKSIWTGT LIALVIYVLW QTAIQXNLPR
251 NEFAPVIAAE GQVSVXIETL SKFAQTGNMD KILSLFSYMA IATSFLGVTL
301 GLFDYIADIF KWNDSVSGRT KTAALTFLPP XISCLLFPTG FVTAIGYVGL
351 AATVWTGIIP AMLLYRSRKK FGAGKTYKVY GGLWLMVWVF LFGIXNIAAX
401 VLSQMELVPV FKG*
```

m102 / a102 95.9% identity in 413 aa overlap

	10	20	30	40	50	. 60
m102.pep	MPNKTPSLFGGAM:	IIAGTVIGAG	MLANPTATSG	VWFTGSLAVL	TVTWECMIC	
				1111111111	131111111	
<b>a</b> 102	MPTKTPSLFGGAMI	IAGTXIGAG	MLANPTATSG	WFTGSLAVI.	LYTWESMISS	CIMITEN
	10	20	30	40	50	60
					30	90
	70	80	90	100	110	120
m102.pep	NTHYPHGASFDTMV	KDLLGRGWN	IINGIAVAFVI	YIJTYAYTE	VCCDI TARCI	CETACCE
		11111111	1111111111		11111111111	111111.
a102	NTHYPHGAXFDTMV	KDLLGRSWN	IINGIAVAFVI	ϓΙΙΤΥΔΥΤΕ	UCCDI MARCI	1111111
	70	80	90	100	110	
				100	110	120
	130	140	150	160	170	180
m102.pep	VSLTVGQLVFFGIL	AFCVWASAR:	LVDRFTGVLIG	GMVT.TETWA	ACCI TADAVA	TOO
		11111111	1	1111111		11111
a102	VSLTVGQLVFFGIL	AFCVWASAR:	LVDRFTSVLIG	CMVITETWA	PCCITADAVI	
	130	140	150	160	170	
				100	170	180
	190	200	210	220	230	240
m102.pep	APAGTNYWIYAATA	LPVCLASFGI	FHGNVSSLLKY	FKGDAPKVA	230 797111122	240
		1 [ ]   1 ]   1   1	1	1111111111		
a102	APTGTNYWIYVATA	LPVCLASFG	HGNVSSLLKY	FKCDAPKUAL		
	190	200	210	220	230	
				220	230	240
	250	260	270	280	290	200
m102.pep	QTAIQGNLPRNEFA	PVIAAEGOVS	VLIETLSKFA	OTGNMDKITS	TECVMATAM	300
		1	1 111111	1111111111	11111111	
a102	QTAIQXNLPRNEFA	PVIAAEGOVS	VXIETLSKFA	OTGNMOKTI 9		
	250	260	270	280	290	
			•	200	230	300
	310	320	330	340	350	360
m102.pep	GLFDYIADIFKWNDS	SISGRTKTAA	LTFLPPLISC	LLFPTGFVTA	TGVVGLAATT	TIMOTTO
		;		1111111111	1111111111	
a102	GLFDYIADIFKWNDS	VSGRTKTAA	LTFLPPXISCI	ʹͺͳͺϜʹϷͲʹϹϝʹͶͲϪ	TGVVCINAMI	MINCTED.
	310	320	330	340	350	360
				510	330	360
	370	380	390	400	410	
m102.pep	AMLLYRSRKKFGAGK	TYKVYGGLW	LMVWVFLFGIV	NIAAOVISO	MEINDABRGA	•
	- 11111111111111111		11111111	1111 1111	1111111111	
<b>a10</b> 2	AMLLYRSRKKFGAGK	TYKVYGGLW	LMVWVFLFCTX	ΝΤΑΔΧΥΙΙΙΙ	-	
370	380 390	400 41	0		AETIA E A E KGY	•
		.50 71	•			

# The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 363>: g105.seq

1	Atgtccgcag	aaaCATACAc	acAAAtcggc	tGGgtaggct	taggGcaaat
51	gggtctgcct	atgGTAACGC	GGCTCTTGGA	CGGCGGCATC	GAAGTCGGCG
101	TATACAACCG	CTCGCCCGAC	AAAACTGCCC	CCATCTCccc	CAAAGGAGCA
151	AAAGTTTACG	GCagcACCGC	CGAACTCGTC	CGCGCCTGCC	$CCCTC\DeltaTTTTT$
201	CCTGATGGTT	TCCGACTATG	CCGCCGTGTG	CGACATCCTG	AACGGAGTCC
251	GCGACGGATT	GGCCGGCAAA	ATCATCGTCA	ACATGAGCAC	CATCTCCCCG
301	ACCGAAAACC	TCGCCGTCAA	AGCACTTGTC	GAAGCCGCAC	GCGGACAGTT
351	TGCCGAAGCA	CCCGTTTCCG	GATCGGTCGG	ACCCCCCACC	AACGGCACAC
401	TGCTGATTCT	GTTCGGCGGC	AGCGAAGCCG	TTTTTAAACCC	GCTGCAAAAA
451	ATATTTTCCC	TTGTCGGCAA	AAAAACCTTC	CATTTCCCCC	ATGTCGGCAA
501	AGGCTCGGGC	GCGAAACTCG	TCTTGAACTC	CATTICGGCG	ATGTCGGCAA
551	AAGCGTACAG	CGAAGCGATG	CTGATGGCGC	GCTCTTAGGC	ATTTTCGGCG
601	GACACCATCG	TCGAACCGAIG	CIGAIGGCGC	GGCAGTTCGG	CATCGATACC
651	TCNANCANAN	1 CGAAGCCA1	CGGCGGCTCG	GCAATGGACT	CGCCTATGTT
701	CAMACAAAA	AAATCACTAT	GGGCAAACCG	TGAGTTCCCC	CCTGCCTTTG
751	CACTCAAACA	CGCTTCCAAA	GACCETAACC	TCGccgtcAA	AGAGCTTGAA
	CAGGCAGGCA	ACACCCTGCC	CGCCGTCGAA	ACCGTTGCTG	CCAGCTACCC
801	CAAAGCAGTT	GAAGCCGGCT	ACGGCGAACA	GGACGTTTCC	GGCGTTTACC
851	TGAAATTGGC	AGAACACTGA			

```
This corresponds to the amino acid sequence <SEQ ID 364; ORF 105.ng>:
     g105.pep
              MSAETYTQIG WVGLGQMGLP MVTRLLDGGI EVGVYNRSPD KTAPISAKGA
          51
              KVYGSTAELV RACPVIFLMV SDYAAVCDIL NGVRDGLAGK IIVNMSTISP
              TENLAVKALV EAAGGQFAEA PVSGSVGPAT NGTLLILFGG SEAVLNPLQK
         101
         151 IFSLVGKKTF HFGDVGKGSG AKLVLNSLLG IFGEAYSEAM LMARQFGIDT
         201 DTIVEAIGGS AMDSPMFQTK KSLWANREFP PAFALKHASK DLNLAVKELE
              QAGNTLPAVE TVAASYRKAV EAGYGEQDVS GVYLKLAEH
The following partial DNA sequence was identified in N. meningitidis <SEO ID 365>:
     m105.seq
              ATGTCCGCAA ACGAATACGC ACAAATCGGC TGGaTAGGCT TAGGGCAAAT
          51
              GGGTCTGCCT ATGGTAACGC GGCTCTTGGA CGGCGCATC GAAGTCGGCG
         101
              TATACAACCG CTCGCCCGAC AAAACTGCCC CCATCTCCGC CAAAGGCGCA
         151 AAAGTTTACG GCAACACCGC CGAACTCGTC CGCGACTATC CCGTCATTTT
         201 CCTGATGGTT TCCGACTATG CCGCCGTGTG CGACATCCTG AACGGAGTCC
         251 GCGACGGATT GGCCGGCAAM ATCATCGTCA ACATGAGCAC CATCTCCCCG
         301
              ACCGAAAaGC TCGCCGTCAA AGCACTTGTC GAAGCGCAGm GaCAGTTTGC
              CGAAGCACCC GTTTCCGGAT CGGTCGGGCC CGCCACCAAC GGCACGCTGC
         351
              TGATTCTGTT CGGCGGCAGC GAACCGLTTT AAACCCGCTG CAAAAAATAT
         401
              TTTCCCTCGT CGGCAAAAA ACCTTCCATT TCGGCGATGT CGGCAAAGGT
         451
              TCGGGCGCGA AACTCGTCTT GAACTCGCTC TTGGGCATTT TCGGCGAaCG
         501
              TACAGCGAAS GmTgCTGATG GCGCGGCAGT TCGGCATCGA TACCGACACC
         551
         601 ATCGTCGAAG CCATCGGSGA CTCGGCAATG GACTCGCCCA TGTTCCAAAC
         651 CAAAAAATCC CTGTGGGCAA ACCGCGAATT CCCGmCCGmC TTCGCCCTCA
         701 AACACGCCTC CAAAGACCTC AACCTCGCCG TCAAAGAGCT TGAACAGGCA
         751 GGCAACACCC TGCCCGCCGT CGAAACCGTT GCTGCCAGCT ACCGCAAAGC
         801 AGTCGAAGCC GGCTACGGGA CACAGGACGT TTCCGGCGTT TACCTGAAAC
              TGGCAGAACA CTGA
         851
This corresponds to the amino acid sequence <SEQ ID 366; ORF 105>:
    m105.pep
              MSANEYAQIG WIGLGOMGLP MVTRLLDGGI EVGVYNRSPD KTAPISAKGA
           1
              KVYGNTAELV RDYPVIFLMV SDYAAVCDIL NGVRDGLAGX IIVNMSTISP
              TEKLAVKALV EAQRQFAEAP VSGSVGPATN GTLLILFGGS EPFXTRCKKY
         151 FPSSAKKPSI SAMSAKVRAR NSSXTRSWAF SANVQRXXLM ARQFGIDTDT
             IVEAIGDSAM DSPMFQTKKS LWANREFPXX FALKHASKDL NLAVKELEQA
              GNTLPAVETV AASYRKAVEA GYGTQDVSGV YLKLAEH
Computer analysis of this amino acid sequence gave the following results:
Homology with a predicted ORF from N. gonorrhoeae
ORF 105 shows 79.9% identity over a 289 as overlap with a predicted ORF (ORF 105.ng)
from N. gonorrhoeae:
    m105/g105
                         10
                                   20
                                            30
                                                      40
    g105.pep
                 MSAETYTQIGWVGLGQMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGSTAELV
                 m105
                 MSANEYAQIGWIGLGQMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAELV
                         10
                                   20
                                            30
                                                      40
                                                               50
                                                                         60
                         70
                                   80
                                            90
                                                     100
                                                              110
     g105.pep
                 RACPVIFLMVSDYAAVCDILNGVRDGLAGKIIVNMSTISPTENLAVKALVEAAGGOFAEA
                    RDYPVIFLMVSDYAAVCDILNGVRDGLAGXIIVNMSTISPTEKLAVKALVEAQR-QFAEA
     m105
                         70
                                   80
                                            90
                                                     100
                                                              110
                        130
                                  140
                                           150
                                                     160
                                                              170
                                                                        180
     g105.pep
                 PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFHFGDVGKGSGAKLVLNSLLG
                 :: ::
                                                                 1:
```

PVSGSVGPATNGTLLILFGGSEPFXTRCKKYFPSSAKKP-SISAMSAKVRARNSSXTRSW

160

170

1.50

BNSDOCID: <WO___9957280A2_I_>

m105

120

130

140

WO 99/57280



		190	200	210	220	230	
g105.pep	P IFGEA	YSEAMLMARQ:	FGIDTDTIV	EAIGGSAMDS	PMFOTKKSL	WANDEEDDAT	240 Alkhask
m105	J •	• • • • • • • • • • • • • • • • • • • •	1 1 1 1 1 1 1 1 1	1111   1111			111111
203	180	VQRXXLMARQI 190	FGIDTDTIV	EAIGDSAMDS			ALKHASK
		200	200	210	220	230	
		250	260	270	280	289	
g105.pep		VKELEQAGNTI	PAVETVAA	SYRKAVEAGY	GEQDVSGVYI	KT A DU	
m105	IIIII DINIA	/KELEOAGNET			<u>                                     </u>	1111	
	240	VKELEQAGNTI 250	260	YRKAVEAGY 270	GTQDVSGVYI 280	KLAEH	
<b>571</b> 0.44 .							
The following	partial DNA	sequence v	was identi	fied in N. n	neningitidis	SSEO ID	367>.
aros.seq							5077.
1 51	ATGTCCGCA	A ACGAATAC	AC ACAAAT	CGGC TGGA	PAGGCT TAG	GGCAAAT	
101	TATACAACC	T ATGGTAAC	AC AAAACT	TGGA CGGC	GGCATC GAA	GTCGGCG	
151	MANGITIAC	G GCAACACC	GC CGAACT	CGTC CGCG1	מתעשים/		
201 251	CCIGAIGGI	1 ICCGACTA	TG CCGCCG	דהדה המארז	TOOMS NAS	~~~~	
301	GCGACGGAI	1 GGCCGGCA	AA ATCATC	GTCA ልሮልጥሪ	27CC7C C7M	CTT C C C C C	
351	I GCCGAAGC	C TCGCCGTC	CG GATCGG	TOGG GOOGG	ברכת כר תחת	200200	
401	IGCIGATIC	I GITUGGUG	GC AGCGAA	מיתיתית ממממ	AACCC CCM		
451 501	MIMITITUO		AA AAAAAC	בישיתי עם טינהים	CCCCC AMO		
551	AGG11CGGG	C GCGAAACT	JG TCTTGA	אכידים הכידיכים	TCCCC AMM	TIM COCCO	
601	GACACCATC	3 TUGAAGCCA	AT CGGCGG	שמעמטרים בייש	CCACE CCC		
651	CCVVVCCVVV	AAATUUUT	ST GGGCAA	שממט ארכב	TOTON COOK	COMMO	
701 751	CCCICAMAC	A COCCIOCAL	AA GACCTC	ארר שרככר	このアスカー カーカイ		
801	CAAAGCAGT	A ACACCCTGC C GAAGCCGGC	CT ACGCCGT	CGAA ACCGT	TGCTG CCAC	CTACCG	
851	T COMMAT I LITTLE	. AL-AACACTC	. Δ				
This correspond	ls to the ami	no acid sequ	uence <sf< th=""><th>Q ID 368:</th><th>ORF 105.</th><th>a&gt;·</th><th></th></sf<>	Q ID 368:	ORF 105.	a>·	
arus.pep							
1 51	MSANEYTQIC	WIGLGOMGI	P MVTRLLE	GGI EVGVY	NRSPD KTAP	ISAKGA	
101	KVYGNTAELV TENLAVKALV	KDIEATETIN	IV SDYAAVC	DII. NGVPD	こしかんひ エチもかい	MOMTON	
151	TIPPAGUUIT	TEGDVGKGS	G AKIVINS	T.T.G. TEGERY	JOEDAN TAKEN	0000	
201	DITARUTGGO	WINDS BIME O.T.	K KSLWANR	EFP PARALI	CUNCE DINIT	QFGIDT AVKELE	
251	QAGNTLPAVE	TVAASYRKA	V EAGYGEQ	DVS GVYLKI	LAEH*		
m105/a105	96 58 33.						
	96.5% ide	ntity in ;	289 aa o	verlap			
		10	20	30	40	50	
m105.pep	MSANEYA	QIGWIGLGQM	GLPMVTRLL	DECTEVEVY	D C D D V M R D T		60 77.1.74.471
a105							
	11011111111111	QIGWIGLGQM	20	DGGIEVGVYN 30	RSPDKTAPI:	SAKGAKVYGN	TAELV
				50	40	50	60
m105.pep	POVEUTE	70	80	90	100	110	119
roo.pcp	I I I I I I I I	LMVSDYAAVCI	DILNGVRDG	LAGKIIVNMS	TISPTENLA	/KALVEAAG-	
a105	RDYPVIFI		DILNGVRDGI	LAGKITVNMS			1111
		70	80	90	100 1ENLA	110	QFAEA 120
	120	130	140	150			120
m105.pep	PVSGSVGF	ATNGTLLILE	GGSEAVINE	150 LOKIESLVC	160 KKTEUECDUC	170	179
a 1 0 E							
a105			CODEMATIVE	LOVIE STAGE	KKTFHFGDVG	KGSGAKLVL	NSLLG
	1	.50 1	40	150	160	170	180
	180	190	200	210	220	230	
					220	230	

```
IFGDV-QRXMLMARQFGIDTDTIVEAIGDSAMDSPMFOTKKSLWANREFPXAFALKHASK
m105.pep
          IFGEAYSEAMLMARQFGIDTDTIVEAIGGSAMDSPMFQTKKSLWANREFPPAFALKHASK
a105
               190
                      200
                              210
                                      220
                                             230
         240
                250
                        260
                               270
                                       280
          DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAEHX
m105.pep
          a105
          DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAEHX
                      260
                              270
                                     280
                                             290
```

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 369>: g105-1.seq

```
ATGTCCGCAG AAACATACAC ACAAATCGGC TGGGTAGGCT TAGGGCAAAT
  1
 51
     GGGTCTGCCT ATGGTAACGC GGCTCTTGGA CGGCGGCATC GAAGTCGGCG
     TATACAACCG CTCGCCCGAC AAAACTGCCC CCATCTCCGC CAAAGGAGCA
101
151 AAAGTTTACG GCAGCACCGC CGAACTCGTC CGCGCCTGCC CCGTCATTTT
    CCTGATGGTT TCCGACTATG CCGCCGTGTG CGACATCCTG AACGGAGTCC
201
251
    GCGACGGATT GGCCGGCAAA ATCATCGTCA ACATGAGCAC CATCTCCCCG
301 ACCGAAAACC TCGCCGTCAA AGCACTTGTC GAAGCCGCAG GCGGACAGTT
351
     TGCCGAAGCA CCCGTTTCCG GATCGGTCGG ACCCGCCACC AACGGCACAC
    TGCTGATTCT GTTCGGCGGC AGCGAAGCCG TTTTAAACCC GCTGCAAAAA
401
451
    ATATTTTCCC TTGTCGGCAA AAAAACCTTC CATTTCGGCG ATGTCGGCAA
     AGGCTCGGGC GCGAAACTCG TCTTGAACTC GCTCTTAGGC ATTTTCGGCG
    AAGCGTACAG CGAAGCGATG CTGATGGCGC GGCAGTTCGG CATCGATACC
551
    GACACCATCG TCGAAGCCAT CGGCGGCTCG GCAATGGACT CGCCTATGTT
601
    TCAAACAAAA AAATCACTAT GGGCAAACCG TGAGTTCCCC CCTGCCTTTG
651
    CACTCAAACA CGCTTCCAAA GACCTTAACC TCGCCGTCAA AGAGCTTGAA
751
    CAGGCAGGCA ACACCCTGCC CGCCGTCGAA ACCGTTGCTG CCAGCTACCG
801 CAAAGCAGTT GAAGCCGGCT ACGGCGAACA GGACGTTTCC GGCGTTTACC
851 TGAAATTGGC AGAACACTGA
```

# This corresponds to the amino acid sequence <SEQ ID 370; ORF 105-1.ng>: g105-1.pep

```
1 MSAETYTQIG WVGLGQMGLP MVTRLLDGGI EVGVYNRSPD KTAPISAKGA
51 KVYGSTAELV RACPVIFLMV SDYAAVCDIL NGVRDGLAGK IIVNMSTISP
101 TENLAVKALV EAAGGQFAEA PVSGSVGPAT NGTLLILFGG SEAVLNPLQK
151 IFSLVGKKTF HFGDVGKGSG AKLVLNSLLG IFGEAYSEAM LMARQFGIDT
201 DTIVEAIGGS AMDSPMFQTK KSLWANREFP PAFALKHASK DLNLAVKELE
251 QAGNTLPAVE TVAASYRKAV EAGYGEQDVS GVYLKLAEH*
```

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 371>: m105-1.seq

```
1 ATGTCCGCAA ACGAATACGC ACAAATCGGC TGGATAGGCT TAGGGCAAAT
 51 GGGTCTGCCT ATGGTAACGC GGCTCTTGGA CGGCGGCATC GAAGTCGGCG
    TATACAACCG CTCGCCCGAC AAAACTGCCC CCATCTCCGC CAAAGGCGCA
101
151 AAAGTTTACG GCAACACCGC CGAACTCGTC CGCGACTATC CCGTCATTTT
201 CCTGATGGTT TCCGACTATG CCGCCGTGTG CGACATCCTG AACGGAGTCC
    GCGACGGATT GGCCGGCAAA ATCATCGTCA ACATGAGCAC CATCTCCCCG
301 ACCGAAAACC TCGCCGTCAA AGCACTTGTC GAAGCCGCAG GCGGACAGTT
    TGCCGAAGCA CCCGTTTCCG GATCGGTCGG GCCCGCCACC AACGGCACGC
351
    TGCTGATTCT GTTCGGCGGC AGCGAAGCCG TTTTAAACCC GCTGCAAAAA
401
    ATATTTCCC TCGTCGGCAA AAAAACCTTC CATTTCGGCG ATGTCGGCAA
501
    AGGTTCGGGC GCGAAACTCG TCTTGAACTC GCTCTTGGGC ATTTTCGGCG
    AAGCGTACAG CGAAnCGATG CTGATGGCGC GGCAGTTCGG CATCGATACC
551
601
    GACACCATCG TCGAAGCCAT CGGsGACTCG GCAATGGACT CGCCCATGTT
    CCAAACCAAA AAATCCCTGT GGGCAAACCG CGAATTCCCG CCCGCCTTCG
651
    CCCTCAAACA CGCCTCCAAA GACCTCAACC TCGCCGTCAA AGAGCTTGAA
    CAGGCAGGCA ACACCCTGCC CGCCGTCGAA ACCGTTGCTG CCAGCTACCG
751
    CAAAGCAGTC GAAGCCGGCT ACGGCGAACA GGACGTTTCC GGCGTTTACC
801
851 TGAAACTGGC AGAACACTGA
```

# This corresponds to the amino acid sequence <SEQ ID 372; ORF 105-1>: m105-1.pep

- 1 MSANEYAQIG WIGLGQMGLP MVTRLLDGGI EVGVYNRSPD KTAPISAKGA
- 51 KVYGNTAELV RDYPVIFLMV SDYAAVCDIL NGVRDGLAGK IIVNMSTISP

```
101 TENLAVKALV EAAGGQFAEA PVSGSVGPAT NGTLLILFGG SEAVLNPLOK
           IFSLVGKKTF HFGDVGKGSG AKLVLNSLLG IFGEAYSEXM LMARQFGIDT
      151
      201 DTIVEAIGDS AMDSPMFQTK KSLWANREFP PAFALKHASK DLNLAVKELE
      251 QAGNTLPAVE TVAASYRKAV EAGYGEQDVS GVYLKLAEH*
  m105-1/g105-1 96.9% identity in 289 aa overlap
                               20
                                                           50
 m105-1.pep
              {\tt MSANEYAQIGWIGLGQMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAELV}
              m: Eineminaanmannaanmannaanem
 g105-1
              MSAETYTQIGWVGLGQMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGSTAELV
                               20
                                        30
                                                 40
                     70
                               80
                                        90
                                                100
                                                                   120
              \verb"RDYPVIFLMVSDYAAVCDILINGVRDGLAGKIIVNMSTISPTENLAVKALVEAAGGQFAEA"
 m105-1.pep
                RACPVIFLMVSDYAAVCDILNGVRDGLAGKIIVNMSTISPTENLAVKALVEAAGGQFAEA
 q105-1
                                        90
                                                100
                                                                   120
                    1.30
                              140
                                       150
                                                160
                                                          170
              PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFHFGDVGKGSGAKLVLNSLLG
 m105-1.pep
              g105-1
              PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFHFGDVGKGSGAKLVLNSLLG
                    130
                             140
                                       150
                                                160
                                                         170
                             200
                                       210
                                                220
 m105-1.pep
             {\tt IFGEAYSEXMLMARQFGIDTDTIVEAIGDSAMDSPMFQTKKSLWANREFPPAFALKHASK}
             g105-1
             IFGEAYSEAMLMARQFGIDTDTIVEAIGGSAMDSPMFQTKKSLWANREFPPAFALKHASK
                    190
                             200
                                      210
                                                220
                                                         230
                    250
                             260
                                       270
                                                280
 m105-1.pep
             DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAEHX
             g105-1
             DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAEHX
                             260
                                      270
                                                280
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 373>:
a105-1.seq
       1 ATGTCCGCAA ACGAATACAC ACAAATCGGC TGGATAGGCT TAGGGCAAAT
      51
          GGGTCTGCCT ATGGTAACGC GGCTCTTGGA CGGCGGCATC GAAGTCGGCG
     101
          TATACAACCG CTCGCCCGAC AAAACTGCCC CCATCTCCGC CAAAGGCGCA
     151 AAAGTTTACG GCAACACCGC CGAACTCGTC CGCGACTATC CCGTCATTTT
     201 CCTGATGGTT TCCGACTATG CCGCCGTGTG CGACATCCTG AACGGAGTCC
          GCGACGGATT GGCCGGCAAA ATCATCGTCA ACATGAGCAC CATCTCCCCG
     251
     301
         ACCGAAAACC TCGCCGTCAA AGCACTTGTC GAAGCCGCAG GCGGACAGTT
     351
          TGCCGAAGCA CCCGTTTCCG GATCGGTCGG GCCCGCCACC AACGGCACGC
         TGCTGATTCT GTTCGGCGGC AGCGAAGCCG TTTTAAACCC GCTGCAAAAA
     401
         ATATTTTCCC TCGTCGGCAA AAAAACCTTC CATTTCGGCG ATGTCGGCAA AGGTTCGGCG GCGAAACTCG TCTTGAACTC GCTCTTGGGC ATTTTCGGCG
     451
     501
     551
         AAGCGTACAG CGAAGCGATG CTGATGGCGC GGCAGTTCGG CATCGATACC
         GACACCATCG TCGAAGCCAT CGGCGGCTCG GCAATGGACT CGCCCATGTT
     601
         CCAAACCAAA AAATCCCTGT GGGCAAACCG CGAATTCCCA CCCGCCTTCG
     651
          CCCTCAAACA CGCCTCCAAA GACCTCAACC TCGCCGTCAA AGAGCTTGAA
     701
          CAGGCAGGCA ACACCCTGCC CGCCGTCGAA ACCGTTGCTG CCAGCTACCG
     801
         CAAAGCAGTC GAAGCCGGCT ACGGCGAACA GGACGTTTCC GGCGTTTACC
         TGAAATTGGC AGAACACTGA
This corresponds to the amino acid sequence <SEQ ID 374; ORF 105-1.a>:
a105-1.pep
         MSANEYTQIG WIGLGQMGLP MVTRLLDGGI EVGVYNRSPD KTAPISAKGA
         KVYGNTAELV RDYPVIFLMV SDYAAVCDIL NGVRDGLAGK IIVNMSTISP
TENLAVKALV EAAGGQFAEA PVSGSVGPAT NGTLLILFGG SEAVLNPLQK
      51
     151 IFSLVGKKTF HFGDVGKGSG AKLVLNSLLG IFGEAYSEAM LMARQFGIDT
         DTIVEAIGGS AMDSPMFQTK KSLWANREFP PAFALKHASK DLNLAVKELE
     201
     251
         QAGNTLPAVE TVAASYRKAV EAGYGEQDVS GVYLKLAEH*
a105-1/m105-1
                99.0% identity in 289 aa overlap
                             20
                                      30
                                                40
                                                         50
            MSANEYTQIGWIGLGQMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAELV
a105-1.pep
```

m105-1	MSANEYAQIGWIGL	_				
	10	20	30	40	50	60
	70	80	90	100	110	120
a105-1.pep	RDYPVIFLMVSDYA	AVCDILNGVE	RDGLAGKIIVN	MSTISPTENL	AVKALVEAA	GGQFAEA
	111111111111111	11111111111	THRIBIT	1111111111	111111111	111111
m105-1	RDYPVIFLMVSDYA	AVCDILNGVE	RDGLAGKIIVN	MSTISPTENL	AVKALVEAA	GGQFAEA
	70	80	90	100	110	120
	130	140	150	160	170	180
a105-1.pep	PVSGSVGPATNGTL	LILFGGSEAV	LNPLQKIFSL	VGKKTFHFGE	VGKGSGAKL	VLNSLLG
		1111111111	111111111			
m105-1	PVSGSVGPATNGTL		-			
	130	140	150	160	170	180
	190	200	210	220	230	240
a105-1.pep	IFGEAYSEAMLMAR	_			ANREFPPAF	
		1111111111	1111 11111		111111111	
m105-1	IFGEAYSEXMLMAR					
	190	200	210	220	230	240
	250	260	270	200	200	
105 1	250	260	270	280	290	
a105-1.pep	DLNLAVKELEQAGN	TLPAVETVA	SYRKAVEAGY	GEQDVSGVYL	KLAEHX	
105 1		1111111111	1111111111	11111111111	111111	
m105-1	DLNLAVKELEQAGN 250	7LPAVETVAA 260	270	GEQDVSGVYL 280		
	250	200	210	280	290	

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 375>:

```
g107.seq
       1 ATGGTATTAA CCTTTATTTG GGCAACCGGC GGCCTGGTTG CCGATGCCAA
      51 ACCGTCCGTC CTCTTCGACA CCCAAGCCCC CGTCGGCACC GGCTACTGGA
     101 TTTACGCCGC CACCGCCCTG CCCGTCTGCC TCGCTTCCTT CGGCTTCCAC
     151 GGCAACGTTT CCAGCCTGCT CAAATACTTT AAAGGCGACG cgcCcaaagt
     201 GgCGAaATCc atctggGcag gtacattggT TGCCttggta atttacgtcc
     251 TCTggcaaac cgccatCcaa agcaaccTGC cgcgcaacga gttcgcCCCc
     301 gtgattgccg ccgagaggca actCTCCGTC CTgaatgaaa cccTGtccaa 351 attcgcccaa accggcgata tggataAaat attqtcccta tttccctaca
     401 tggcaatcgc cacctccttt ttaggcgTAA Ccttaggcct gtttgacaac
     451 atcgccggac atcttcaaat ggaacgacag tatgtccggg cggcaccaaa
     501 accgtcgcgc tga
```

This corresponds to the amino acid sequence <SEO ID 376; ORF 107.ng>: q107.pep

- MVLTFIWATG GLVADAKPSV LFDTQAPVGT GYWIYAATAL PVCLASFGFH
- 51 GNVSSLLKYF KGDAPKVAKS IWAGTLVALV IYVLWQTAIQ SNLPRNEFAP 101 VIAAERQLSV LNETLSKFAQ TGDMDKILSL FPYMAIATSF LGVTLGLFDN
- 151 IAGHLQMERQ YVRAAPKPSR *

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 377>: m107.seq

- 1 ATGGTATTGA CCTTTATTTG GGCGGCCGGC GGGCTGATTG CCGATGCCAA 51 GCCGTCCGTC CTCTTCGATA CCCAAGCCCC CGCCGGCACA AACTACTGGA
- 101 TTTACGCCGs CACCGCCCTG CCCGTCTGCC TCGCTTCCTT CGGCTTCCAC
- 151 GGCAACGTCT CCAGCCTGCT CAAATACTTT AAAGGCGACG CGCCCAAAGT
- 201 GGCTAAATCC ATCTGGACGG GCACACTGAT TGCGCTGGTA ATTTACGTCC
- 251 TCTGGCAAAC CGCCATCCAA GGCAACCTGC CGCGCAACGA GTTCGCCCCC
- 301 GTCATCGCCG CCGAAGGGCA AGTCTCCGTC CTCATCGAAA CCCTGTCCAA
- 351 ATTCGCCCAA ACCGGCAATA TGGACAAAAT ATTGTCCCTG TTTTCCTATA
- 401 TGGCGATCGC CACCTCGTTT TTAGGCGTAA CGCTCGGACT CTTCGACTAC
- 451 ATCGCCCATC TTCAAATGGA ACGACAGCAT CTCCGGqCCG CACCAAAACC
- 501 GCCGCGCTGA

This corresponds to the amino acid sequence <SEQ ID 378; ORF 107>: m107.pep..

- 1 MVLTFIWAAG GLIADAKPSV LFDTQAPAGT NYWIYAXTAL PVCLASFGFH
- 51 GNVSSLLKYF KGDAPKVAKS IWTGTLIALV IYVLWQTAIQ GNLPRNEFAP

WO 99/57280



101 VIAAEGQVSV LIETLSKFAQ TGNMDKILSL FSYMAIATSF LGVTLGLFDY IAHLQMERQH LRAAPKPPR* Computer analysis of this amino acid sequence gave the following results: Homology with a predicted ORF from N. gonorrhoeae ORF 107 shows 89.4% identity over a 170 aa overlap with a predicted ORF (ORF 107.ng) from N. gonorrhoeae: m107/q107 10 20 30 40 60 t MVLTFIWAAGGLIADAKPSVLFDTQAPAGTNYWIYAXTALPVCLASFGFHGNVSSLLKYFm107.pep MVLTFIWATGGLVADAKPSVLFDTQAPVGTGYWIYAATALPVCLASFGFHGNVSSLLKYF 9107 10 20 30 40 50 70 80 90 100 110 KGDAPKVAKSIWTGTLIALVIYVLWQTAIQGNLPRNEFAPVIAAEGQVSVLIETLSKFAQ 120 m107.pep KGDAPKVAKSIWAGTLVALVIYVLWQTAIQSNLPRNEFAPVIAAERQLSVLNETLSKFAQ q107 70 80 90 100 110 120 130 140 150 160 170 TGNMDKILSLFSYMAIATSFLGVTLGLFDYIA-HLQMERQHLRAAPKPPR m107.pep q107 TGDMDKILSLFPYMAIATSFLGVTLGLFDNIAGHLQMERQYVRAAPKPSR 130 150 160 170 The following partial DNA sequence was identified in N. meningitidis <SEQ ID 379>: a107.seq ATGGTATTAA CCTTTATTTG GGCAACCGGC GGCCTGATTG CCGATGCCAA ACTGCCCGTC CTCTTCGACA CCCAAGCCCC TACCGGCACC AACTACTGGA 51 TTTATGTCGC CACCGCCCTG CCCGTCTGCC TTGCGTCATT CGGTTTCCAC 101 151 GGCAACGTCT CCAGCCTGCT CAAATACTTT AAAGGCGACG CGCCCAAAGT GGCTAAATCC ATCTGGACGG GCACACTGAT TGCGCTGGTA ATTTACGTCC TCTGGCAAAC CGCCATCCAA GGCAACCTGC CGCGCAACGA GTTCGCCCCC 251 GTGATTGCCG CCGAAGGGCA AGTCTCCGTC CTGATTGAAA CCCTGTCCAA 301 ATTCGCCCAA ACCGGCAATA TGGACAAAAT ATTGTCCCTG TTTTCCTATA 351 TGGCGATCGC CACCTCGTTT TTAGGCGTAA CGCTCGGACT CTTCGACTAC 401 ATCGCCGACA TCTTCAAATG GAACGACAGC GTGTCCGGCC GCACCAAAAC CGCCGCGCTG ACCTTCCTGC CGCCTCTAAT TTCCTGCCTG CTCTTCCCCA 501 CCGGCTTTGT TACCGCCATC GGCTACGTCG GCCTGGCGGC AACCGTCTGG 551 ACAGGCATCA TCCCCGCCAT GCTGCTCTAC CGTTCGCGCA AAAAATTCGG 601 651 CGCAGGCAAA ACCTATAAAG TTTACGGCGG CTTGTGGCTG ATGGTTTGGG TCTTCCTTTT CGGCATCGTC AACATCGCCG CACAGGTATT GAGCCAAATG 701 751 GAACTCGTCC CCGTATTTAA AGGATAA This corresponds to the amino acid sequence <SEQ ID 380; ORF 107.a>: a107.pep MVLTFIWATG GLIADAKLPV LFDTQAPTGT NYWIYVATAL PVCLASFGFH GNVSSLLKYF KGDAPKVAKS IWTGTLIALV IYVLWQTAIQ GNLPRNEFAP 51 VIAAEGQVSV LIETLSKFAQ TGNMDKILSL FSYMAIATSF LGVTLGLFDY 101 IADIFKWNDS VSGRTKTAAL TFLPPLISCL LFPTGFVTAI GYVGLAATVW 151 201 TGIIPAMLLY RSRKKFGAGK TYKVYGGLWL MVWVFLFGIV NIAAQVLSQM 251 ELVPVFKG* m107/a107 94.8% identity in 154 aa overlap 20 30 40 MVLTFIWAAGGLIADAKPSVLFDTQAPAGTNYWIYAXTALPVCLASFGFHGNVSSLLKYF m107.pep

THEOREMAIN THOROGODIE THOROGODIENTO MVLTFIWATGGLIADAKLPVLFDTQAPTGTNYWIYVATALPVCLASFGFHGNVSSLLKYF

40

30

20

a107

	70	80	90	100	110	120
m107.pep	KGDAPKVAKSIWTG	TLIALVIYVL	WQTAIQGNLP	RNEFAPVIA	AEGQVSVLIET	LSKFAQ
		111111111	1111111111	1111111111		11111
a107	KGDAPKVAKSIWTG	TLIALVIYVL	WQTAIQGNLP	RNEFAPVIA	AEGQVSVLIE1	LSKFAQ
	70	80	90	100	110	120
	130	140	150	160	170	
m107.pep	TGNMDKILSLFSYM	AIATSFLGVT	LGLFDYIAHL	QMERQHLRAZ	APKPPRX	
	1111111111111	1111111111	11111111:			
a107	TGNMDKILSLFSYM	AIATSFLGVT	LGLFDYIADI	FKWNDSVSGI	RTKTAALTFLE	PLISCL
	130	140	150	160	170	180
			D			
a107	LFPTGFVTAIGYVG					
	190	200	210	220	230	240

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 381>:

```
g108.seq
```

- 1 ATGttgccgg gCTTCAACCG GATATTCAaa cggTTTGCTC CAACACTCGG
  - 51 AACGGCGCAT AAAACGCCGC CCTTCGCGTT ATCCCGAACG GGGCGGCTAA
- 101 TCAGATCCTA TCGCCATAAA AGGCGGGGTT TCAACCGAAA AGGAATTGAG
- 151 ATGAATAAAA CCTTGTCTAT TTTGCCGGCG GCAATCTTAC TCGGCGGGTG
- 201 CGCCGCCGGC GGCAACACAT TCGGCAGCTT AGACGGCGGC ACGGGTATGG
- 251 GTGGCAGCAT CGTCAAAATG ACGGTAGAAA gccAATGCCG TGCGGAATTG
- 301 GACAGGCGCA GCGAATGGCG TTTGACCGCG CTGGCGATGA GTGCCGAAAA
- 351 ACAGGCGGAA TGGGAAAACA AGATTTGCGG CTGCGCTACC GAAGAAGCAC 401 CTAACCAGCT GACCGGCAAC GATGTGATGC AGATGCTGAa ccagtccacG
- 451 CGCaatcagg cacTtgccgc CCtgaccgTC AAAacggtTT CcgcctgcTT
- 501 CAaacgcctg tACCGCTAa

This corresponds to the amino acid sequence <SEQ ID 382; ORF 108.ng>:

g108.pep

- 1 MLPGFNRIFK RFAPTLGTAH KTPPFALSRT GRLIRSYRHK RRGFNRKGIE
- 51 MNKTLSILPA AILLGGCAAG GNTFGSLDGG TGMGGSIVKM TVESQCRAEL
- 101 DRRSEWRLTA LAMSAEKQAE WENKICGCAT EEAPNQLTGN DVMQMLNQST
- 151 RNQALAALTV KTVSACFKRL YR*

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 383>:

m108.seq

- 1 ATGTTGCCGG GCTTCAACCG GATATTCAAA CGGTTTGTTC CAACACTCGG
- 51 AACGGCGCAT AAAACGCCGC CCTTCGCGTT ATCCCGAACG GGGCGGCTAA
- 101 TCAGATTCTA TCGCCATAAA AGGCGGGGTT TCAACCGAAA AGGAATTGAG
- 151 ATGAATAAAA CCTTGTCTAT TTTGCCGGTG GCAATCTTAC TCGGCGGCTG
- 201 CGCCGCCGGA GGCGGTAACA CATTCGGCAG CTTAGACGGT GGCACAGGCA
- 251 TGGGCGGCAG CATCGTCAAA ATGGCGGTTG GGAGCCAATG CCGTGCGGAA
- 301 TTGGACAAAC GCAGCGAATG GCGTTTGACC GCGCTGGCGA TGAGTGCCGA
- 351 AAAACAGGCG GAGTGGGAAA ACAAGATTTG CGCTTGCGTC GCCCAAGAAG
- 401 CACCCGAACG GATGACCGGC AACGATGTGA TGCAGATGCT GGCTCCGTCC
- 451 ACGCGCAATC AGGCACTTGC CGCCCTGACC GCCAAAACGG TTTCCGCCTG
- 501 CTTCAAACAC CTGTACCGCT AA

This corresponds to the amino acid sequence <SEQ ID 384; ORF 108>:

m108.pep

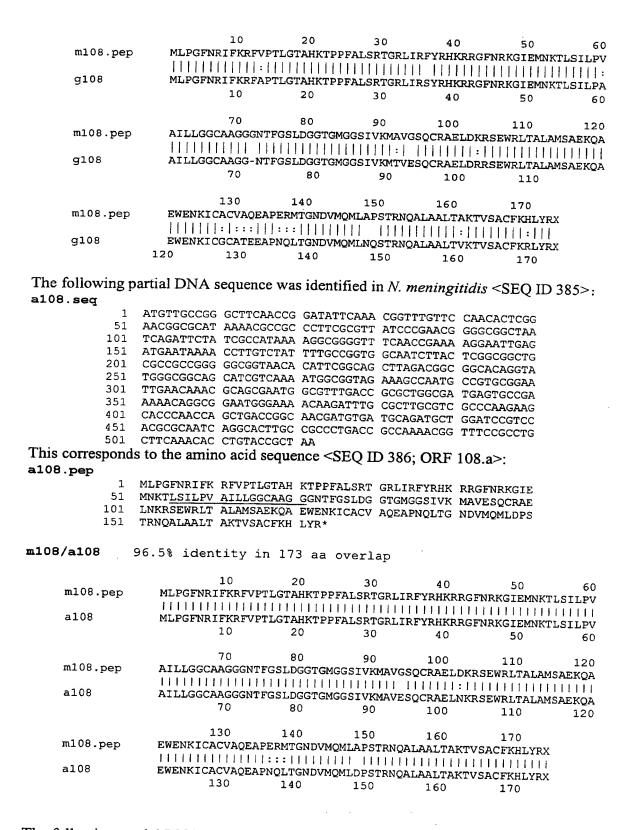
- 1 MLPGFNRIFK RFVPTLGTAH KTPPFALSRT GRLIRFYRHK RRGFNRKGIE
- 51 MNKTLSILPV AILLGGCAAG GGNTFGSLDG GTGMGGSIVK MAVGSQCRAE
- 101 LDKRSEWRLT ALAMSAEKQA EWENKICACV AQEAPERMTG NDVMQMLAPS
- 151 TRNQALAALT AKTVSACFKH LYR*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae

ORF 108 shows 89.6% identity over a 173 as overlap with a predicted ORF (ORF 108.ng) from N. gonorrhoeae:

m108/g108



The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 387>:

```
g109.seq
         ATGTATTATC GCCGGGTTGT GGGGCTATCC GATGGACTTG GCGATTTGGC
      1
      51 AGCCGGTATT GATCGTAGGC GTATGCTTAC CGCTTTTGGA AGCGGGCATG
     101 GAAATGACGC GCAAAGGCAA AACCACCCAA TCCGCCGCCA TCGTGGTGTT
     151 CTCTTCCGTC TGGTCAATCC GGTTTTCGGC TGGGCGTTGA CGATGCTGTT
     201 GGATAATTTG GGCTTAATCG GCTGCAAAGA ACGCAGCGCG CAATTAGGTT
     251 TTGTCGGACG AGTATTGATA CCCGCAGTAG GTTTCTTAAT CTTGTGTGTG
     301 GCGATGGGTG CGGTCGGGAT GCTGCCCGGT ATCCCTCCGT TTTTGGAGCA
     351 GTTCAAATCT TTGGGCTAG
This corresponds to the amino acid sequence <SEQ ID 388; ORF 109.ng>:
g109.pep
         MYYRRVVGLS DGLGDLAAGI DRRRMLTAFG SGHGNDAQRQ NHPIRRHRGV
       1
         LFRLVNPVFG WALTMLLDNL GLIGCKERSA QLGFVGRVLI PAVGFLILCV
      51
         AMGAVGMLPG IPPFLEQFKS LG
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 389>:
m109.seq
         ATGTATTATC GCCGGGTTAT GGGGCTATCC GATGGACTTG GCGATTTGGC
       1
      51 AGCCGGTATT GAGCGTAGCC TTGGTCGTAG GCGTATACTT ACCGCTTTTG
     101 GAAGCGGGCA TGGAAATGAC GCGCAAAGGC AAAACCACCC AATCCGCCGC
     151 CATCGTGGTG TTCTCTTCCG CCTTGTCAAT CCGGTTTTCG GCTGGGCGTT
     201 GACGATGCTG TTGGATAATT TGGGCTTAAT CGGCTGCAAA GAGCGCAGTG
     251 CGCAATTAGG TTTCGCCGGA CGCGTGTTGA TACCCGCAGT AGGTTTCTTG
     301 ATCTTGTGTG TGGCGATGGG TGCGGTCGGG ATGCTGCCCG GTATCCCGCC
     351 GTTTTTGGAA CACTTCAAAT CTTTGGGCTA G
This corresponds to the amino acid sequence <SEQ ID 4; ORF 109>:
m109.pep
          MYYRRVMGLS DGLGDLAAGI ERSLGRRRIL TAFGSGHGND AQRQNHPIRR
      51 HRGVLFRLVN PVFGWALTML LDNLGLIGCK ERSAQLGFAG RVLIPAVGFL
     101 ILCVAMGAVG MLPGIPPFLE HFKSLG*
Computer analysis of this amino acid sequence gave the following results:
Homology with a predicted ORF from N. gonorrhoeae
ORF 109 shows 92.9% identity over a 126 aa overlap with a predicted ORF (ORF 109.ng)
from N. gonorrhoeae:
m109/g109
                              20
                                        30
                                                  40
             MYYRRVMGLSDGLGDLAAGIERSLGRRRILTAFGSGHGNDAQRQNHPIRRHRGVLFRLVN
m109.pep
             MYYRRVVGLSDGLGDLAAGIDR----RRMLTAFGSGHGNDAQRQNHPIRRHRGVLFRLVN
g109
                              20
                                            30
                                                     40
                                        90
                                                 100
                                                          110
                     70
                              80
             {\tt PVFGWALTMLLDNLGLIGCKERSAQLGFAGRVLIPAVGFLILCVAMGAVGMLPGIPPFLE}
m109.pep
             PVFGWALTMLLDNLGLIGCKERSAQLGFVGRVLIPAVGFLILCVAMGAVGMLPGIPPFLE
g109
                        70
                                  80
                                            90
                                                     100
             HFKSLGX
m109.pep
             : | | | | |
g109
             QFKSLGX
              120
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 391>:
a109.seq
              ATGTATTATC GCCGGGTTGT GGGGCTATCC GATGGACTTG GCGATTTGGC
           51 AGCCGGTATT GAGCGTAGCC TTGGTCGTAG GCGTATACTT ACCGCTTTTG
          101 GAAGCGGCA TGGAAATGAC GCGCAAAGGC AAAACCACCC AATCCGCCGC
```

151 CACCGTGGTG TTCTCTCCG CTTGGTCAAT CCGGTTTTCG GCTGGGCGTT

BNSDOCID: <WO___9957280A2_|_>

201	GACGATGCTG	TTGGATAATT	TGGGCTTAAT	CGGCTGCAAA	GAGCGCACCC
251	CGCAATTAGG	TTTCACCGGA	CGCGTATTGA	TACCCGTAGT	ACCUTTUCTUC
301	ATCTTGTGTG	TGGCGATGGG	TGCGGTCGGG	ATGCTGCCCG	CTATCCCCCC
351	GTTTTTGGAG	CACTTCAAAT	CTTTGGGCTA	G	GIMICCCGCC

## This corresponds to the amino acid sequence <SEQ ID 392; ORF 109>: a109.pep

- MYYRRVVGLS DGLGDLAAGI ERSLGRRRIL TAFGSGHGND AQRQNHPIRR
- HRGVLFRLVN PVFGWALTML LDNLGLIGCK ERSAQLGFTG RVLIPVVGFL
- 101 <u>ILCVAMGA</u>VG MLPGIPPFLE HFKSLG*

#### 97.6% identity in 126 aa overlap m109/a109

```
20
                                        40
                                               50
          MYYRRVMGLSDGLGDLAAGIERSLGRRRILTAFGSGHGNDAQRQNHPIRRHRGVLFRLVN
m109.pep
          MYYRRVVGLSDGLGDLAAGIERSLGRRRILTAFGSGHGNDAQRQNHPIRRHRGVLFRLVN
a109
                10
                        20
                                30
                                        40
                                               50
                70
                        80
                                90
                                       100
                                               110
                                                      120
          PVFGWALTMLLDNLGLIGCKERSAQLGFAGRVLIPAVGFLILCVAMGAVGMLPGIPPFLE
m109.pep
          PVFGWALTMLLDNLGLIGCKERSAQLGFTGRVLIPVVGFLILCVAMGAVGMLPGIPPFLE
a109
                        80
                                90
                                      100
                                              110
m109.pep
          HFKSLGX
          111111
a109
          HFKSLGX
```

#### The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 393>: glll.seq

```
ATGCCGTCTG AAACACGCCT GCCGAACCTT ATCCGCGCCT TGATATTTGC
    CCTGGGTTTC ATCTTCCTGA ACGCCTGTTC GGaacaaacC GCGCAaaccg
 51
101 TTACCCTGCA AGGCGAAACG ATGGGTACGA CCLATACCGT CAAATACCTT
151 TCAAATAATC GGGACAAACT CCCCTCCCCT GCCAAAATAC AAAAGCGCAT
201 TGATGATGCG CTTAAAGAAG TCAACCGGCA GATGtccaCC TACCAGACCG
251 ATTCCGAAAT CAGCCGGTTt atacagacan atgctggaga gctcttcgcg
301 tntcatgcag nttctataac tgattccgcc gaagactgtc tgcctaatac
351 gcctatctca tcggcgctct ga
```

## This corresponds to the amino acid sequence <SEQ ID 394; ORF 111.ng>: glll.pep

- MPSETRLPNL IRALIFALGF IFLNACSEQT AQTVTLQGET MGTTYTVKYL SNNRDKLPSP AKIQKRIDDA LKEVNROMST YQTDSEISRF IQTAGELFAH 51
- 101 ASITDSAEDC LPNTPISSAL *

### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 395>: mlll.seg

1	ATGCCGTCTG	AAACACGCCT	GCCGAACTTT	ATCCGCGTCT	TGATATTTGC
51	CCTGGGTTTC	ATCTTCCTGA	ACGCCTGTTC	GGAACAAACC	GCCCAAACCC
101	TTACCCTGCA	AGGCGAAACG	ATGGGCACGA	CCTATAVCCT	CAAATACCTT
151	TCAAATAATC	GGGACAAACT	CCCCTCACCT	GCCGAAATAC	AWAAACGCAT
201	CGATGACGCG	CTTAAAGAAk	TCAACCGGVA	GATGTCCACC	TATCACCCCAI
251	ACTCCGAAAT	CAGCCGGTTC	AACCAACACA	CAGCCGGGAA	CCCCCCCCCC
301	ATTTCAAGCG	ACTTCGCACA	CGTTACTGCC	GAAGCCCTCC	CCCCCTCCGC
351	CCTGACACAC	GGCGCGCTGG	ACGTAACCGT	CGGCCCCTTC	GCC1GAACCG
401	GGGGATTCGG	CCCCGACAAA	TCCGTTACCC	GTGAACCCTC	GICAACCTTT
451	ATCAAACAGG	CGGCATCTTA	TACGGGCATA	GACAAAACCGIC	GCCGGAACAA
501	AGGCAAAGAT	TACGCTTCCT	TGAGCAAAAC	GACAAAATCA	TTTTGAAACA
551	ATTTATCTTC	GATTGCCAAA	CCCTTCCCCC	CCACCCCAAG	GCCTATTTGG
			000110666	TIGATAAAGT	TGCGGGCGAA

```
601
     CTGGAAAAAT ACGGCATTCA AAATTATCTG GTCGAAATCG GCGGCGAGTT
      GCACGGCAAA GGCAAAAACG CGCGCGGCGA ACCGTGGCGC ATCGGTATCG
 651
 701
      AGCAGCCCAA TATCGTCCAA GGCGGCAATA CGCAGATTAT CGTCCCGCTG
     AACAACCGTT CGCTTGCCAC TTCCGGCGAT TACCGTATTT TCCACGTCGA
 751
     TAAAAACGGC AAACGCCTCT CCCATATCAT CAACCCGAAC AACAAACGAC
 801
     CCATCAGCCA CAACCTCGCC TCCATCAGCG TGGTCGCAGA CAGTGCGATG
 851
     ACGGCGGACG GCTTGTCCAC AGGATTATTC GTATTGGGCG AAACCGAAGC
 901
951 CTTAAAGCTG GCAGAGCGCG AAAAACTCGC TGTTTTCCTG ATTGTCAGGG
1001 ATAAAGGCGG CTACCGCACC GCCATGTCTT CCGAATTTGA AAAACTGCTC
1051
     CGCTAA
```

This corresponds to the amino acid sequence <SEQ ID 396; ORF 111>:

```
mlll.pep
```

```
1 MPSETRLPNF IRVLIFALGF IFLNACSEQT AQTVTLQGET MGTTYXVKYL
51 SNNRDKLPSP AEIXKRIDDA LKEXNRXMST YQPDSEISRF NQHTAGKPLR
101 ISSDFAHVTA EAVRLNRLTH GALDVTVGPL VNLWGFGPDK SVTREPSPEQ
151 IKQAASYTGI DKIILKQGKD YASLSKTHPK AYLDLSSIAK GFGVDKVAGE
201 LEKYGIQNYL VEIGGELHGK GKNARGEPWR IGIEQPNIVQ GGNTQTIVPL
251 NNRSLATSGD YRIFHVDKNG KRLSHIINPN NKRPISHNLA SISVVADSAM
301 TADGLSTGLF VLGETEALKL AEREKLAVFL IVRDKGGYRT AMSSEFEKLL
351 R*
```

ORF 111 shows 88.7% identity over a 97 aa overlap with a predicted ORF (ORF 111.ng) from N. gonorrhoeae:

```
m111.pep/g111.pep
```

```
20
                                   30
                                            40
                                                    50
           MPSETRLPNFIRVLIFALGFIFLNACSEQTAQTVTLQGETMGTTYXVKYLSNNRDKLPSP
m111.pep
           MPSETRLPNLIRALIFALGFIFLNACSEQTAQTVTLQGETMGTTYTVKYLSNNRDKLPSP
g111
                                                    50
                  10
                          20
                                   3.0
                                            40
                                                             60
                  70
                           80
                                   90
                                           100
                                                   110
                                                            120
m111.pep
           AEIXKRIDDALKEXNRXMSTYQPDSEISRFNQHTAGKPLRISSDFAHVTAEAVRLNRLTH
           AKIQKRIDDALKEVNRQMSTYQTDSEISRFIQTXAGELFAXHAXSITDSAEDCLPNTPIS
g111
                  70
                          80
                                   90
                                           100
                                                   110
                                                            120
                 130
                          140
                                  150
                                           160
                                                   170
                                                            180
m111.pep
           GALDVTVGPLVNLWGFGPDKSVTREPSPEQIKQAASYTGIDKIILKQGKDYASLSKTHPK
q111
           SALX
```

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 397>: a111.seq

```
ATGCCGTCTG AAACACGCCT GCCGAACTTT ATCCGCACCT TGATATTTGC
51
    CCTGAGTTTT ATCTTCCTGA ACGCCTGTTC GGAACAAACC GCGCAAACCG
101
    TTACCCTGCA AGGTGAAACG ATGGGCACGA CCTATACCGT CAAATACCTT
151
    TCAAATAATC GGGACAAACT CCCCTCACCT GCCGAAATAC AAAAGCGCAT
201
     CGATGACGCG CTTAAAGAAG TCAACCGGCA GATGTCCACC TATCAGCCCG
    ACTCCGAAAT CAGCCGGTTC AACCAACACA CAGCCGGCAA GCCCCTCCGC
251
301
     ATTTCAAGCG ACTTCGCACA CGTTACTGCC GAAGCCGTCC ACCTGAACCG
351
     CCTGACACAC GGCGCGCTGG ACGTAACCGT CGGCCCCTTG GTCAACCTTT
401
     GGGGATTCGG CCCCGACAAA TCCGTTACCC GTGAACCGTC GCCGGAACAA
451
     ATCAAACAAG CAGCATCTTA TACGGGCATA GACAAAATCA TTTTGAAACA
    AGGCAAAGAT TACGCTTCCT TGAGCAAAAC CCACCCCAAG GCCTATTTGG
501
    ATTTATCTTC GATTGCCAAA GGCTTCGGCG TTGATAAAGT TGCGGGCGAA
551
601
    CTGGAAAAAT ACGGCATTCA AAATTATCTG GTCGAAATCG GCGGCGAGTT
651
    GCACGGCAAA GGCAAAAACG CGCGCGGCGA ACCTTGGCGC ATCGGCATCG
701
    AACAGCCCAA CATCGTCCAA GGCGGCAATA CGCAGATTAT CGTCCCGCTG
751
    AACAACCGTT CGCTTGCCAC TTCCGGCGAT TACCGTATTT TCCACGTCGA
```

801 851 901 951 1001 1051	TAAAAGCGGC AAACGCCTCT CCCATATCAT TAATCCGAAC AACAAACGAC CCATCAGCCA CAACCTCGCC TCCATCAGCG TGGTCGCAGA CAGTGCGATG ACGGCGGACG GCTTGTCCAC AGGATTATTC GTATTGGGCG AAACCGAAGC CTTAAAGCTG GCAGAGCGCG AAAAACTCGC TGTTTTCCTG ATTGTCAGGG ATAAAGGCGG CTACCGCACC GCCATGTCTT CCGAATTTGA AAAACTGCTC CGCTAA
This correspond	s to the amino acid sequence <seq 111.a="" 398;="" id="" orf="">:</seq>
all1.pep	o to the annie acid sequence (SEQ ID 396; ORF 111.a>:
errr.beb	
1	MPSETRLPNF IRTLIFALSF IFLNACSEQT AQTVTLQGET MGTTYTVKYL
51	SNNRDKLPSP AEIQKRIDDA LKEVNRQMST YQPDSEISRF NQHTAGKPLR
101	ISSDFAHVTA EAVHLNRLTH GALDVTVGPL VNLWGFGPDK SVTREPSPEO
151	TKOAASYTGI DKIILKOCKD VASISKTURK AND THE STREPSPEQ
201	IKQAASYTGI DKIILKQGKD YASLSKTHPK AYLDLSSIAK GFGVDKVAGE
	LEKYGIQNYL VEIGGELHGK GKNARGEPWR IGIEQPNIVQ GGNTQIIVPL
251	NNRSLATSGD YRIFHVDKSG KRISHIINDN NVDDIGUNIA

251 NNRSLATSGD YRIFHVDKSG KRLSHIINPN NKRPISHNLA SISVVADSAM 301 TADGLSTGLF VLGETEALKL AEREKLAVFL IVRDKGGYRT AMSSEFEKLL

### ml11/al11 97.7% identity in 351 aa overlap

m111.pep	10 MPSETRLPNFIRV             MPSETRLPNFIRT   10			1 1 1 1 1 1 1 1 1		
m111.pep	70	80	90	100	110	120
mrrr.beb	AEIXKRIDDALKEX	NRXMSTYOPI	DSEISRFNOHT	'AGKPLRISS	DFAHVTAEAV.	RLNRLTH
a111	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1			1111111	1111111111	
	AEIQKRIDDALKEV 70	14Ö1 15MÖYN. 80	OSEISKFNOHT 90	AGKPLRISSI		HLNRLTH
	, 0	80	90	100	110	120
	130	140	150	160	170	100
m111.pep	GALDVTVGPLVNLW	GFGPDKSVTF	EPSPEOTKOA	ASYTCTDET	T VOCUDURA	180
		3   1   1   1   1   1	1 1 1 1 1 1 1 1 1 1 1	11111111		
a111	GWTDAIAGETANTM	GEGPDKSVTR	EPSPEQIKQA	ASYTGIDKI	LKOGKDYASI	SKTHPK
	130	140	150	160	170	180
	190	200	010			
mlll.pep			210	220	230	240
* *	AYLDLSSIAKGFGV	LIIIIIIIIII	GIÖNITAFIG(	SELHGKGKNA	RGEPWRIGIE	QPNIVQ
a111	AYLDLSSIAKGFGV	DKVAGELEKY	GTONYLVETG	IIIIIIIIII	DCDD1777	11111
	190	200	210	220	RGEPWRIGIE 230	
			220	220	230	240
- 111	250	260	270	280	290	300
mll1.pep	GGNTQIIVPLNNRS:	LATSGDYRIF	HVDKNGKRLSF	HIINPNNKRP	TCUNIT NOTON	777 5 6 5 5 5
a111	*	3	1111:1111		111111111	
G111	GGNTQIIVPLNNRS	PHISCHIKIE	HVDKSGKRLSE	HIINPNNKRP	ISHNLASISV	VADSAM
	230	260	270	280	290	300
	310	320	330	340		
m111.pep	TADGLSTGLFVLGET		KLAVFLTVRDK	CCVDTAMOO	350	
		1		TITTLEFE	LICEPLE	
a111	TAUGLSTGLFVLGET	EALKLAERE	KLAVFLIVRDK	GGYRTAMSS	EFEKLIUV	
	310	320	330	340	350	

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 399>:

- 1 ATGCCGTCTG AAACACGCCT GCCGAACCTT ATCCGCGCCT TGATATTTGC
  51 CCTGGGTTTC ATCTTCCTGA ACGCCTGTTC GGAACAAACC GCGCAAaCCG
  101 TTACCCTGCA AGGCGAAACG ATGGGTACGA CCTATACCGT CAAATACCTT
  151 TCAAATACTT CCCACAAACG CCCACAAACC
- TCAAATAATC GGGACAAACT CCCCTCCCCT GCCAAAATAC AAAAGCGCAT
  TGATGATGCG CTTAAAGAAG TCAACCGGCA GATGTCCACC TACCAGACCG

```
251 ATTCCGAAAT CAGCCGGTTC AACCAACACA CAGCCGGCAA GCCCCTCCGC
     ATTTCAAGCG ATTTCGCACA CGTTACCGCC GAAGCCGTCC GCCTGAACCG
301
351
     CCTGACTCAC GGCGCACTGG ACGTAACCGT CGGCCCTTTG GTCAACCTTT
401
     GGGGGTTCGG CCCCGACAAA TCCGTTACCC GTGAACCGTC GCCGGAACAA
     ATCAAACAGG CGGCATCTTA TACGGGCATA GACAAAATCA TTTTGCAACA
451
     AGGCAAAGAT TACGCTTCCT TGAGCAAAAC CCACCCAAA GCCTATTTGG
ATTTATCTTC GATTGCCAAA GGCTTCGGCG TTGATAAAGT TGCGGGCGAA
501
551
     CTGGAAAAAT ACGGCATTCA AAATTATCTG GTCGAAATCG GCggcGAGTT
601
     GCACGGCAAA GGCAAAAATG CGCACGGCGA ACCGTGGCGC ATCGGTATAG
651
701
     AGCAACCCAA TATcatccaa ggcggcaata cgcAGattat cgtcccgctg
751
     aaCaaccgtt cgcttgccac ttccggcgAT taccgtaTTT tccacgtcgA
     TAAAAACGGC Aaacgccttt cccacATCAT CAATCCCAAC AACAAACGAC
801
     CCATCAGCCA CAACCTCGCC tcCATCAGCG TGGTCTCAGA CAGTGCAATG
851
901
     ACGGCGGACG GTTTATCCAC AGGATTATTT GTTTTAGGCG AAACCGAAGC
     CTTAAGGCTG GCAGAACAAG AAAAACTCGC TGTTTTCCTA ATTGTCCGGG
951
     ATAAGGACGG CTACCGCACC GCCATGTCTT CCGAATTTGC CAAGCTGCTC
```

This corresponds to the amino acid sequence <SEQ ID 400; ORF 111-1.ng>: g111-1.pep

```
1 MPSETRLPNL IRALIFALGF IFLNACSEQT AQTVTLQGET MGTTYTVKYL
51 SNNRDKLPSP AKIQKRIDDA LKEVNRQMST YQTDSEISRF NQHTAGKPLR
101 ISSDFAHVTA EAVRLNRLTH GALDVTVGFL VNLWGFGPDK SVTREPSPEQ
151 IKQAASYTGI DKIILQQGKD YASLSKTHPK AYLDLSSIAK GFGVDKVAGE
201 LEKYGIQNYL VEIGGELHGK GKNAHGEPWR IGIEQPNIIQ GGNTQIIVPL
251 NNRSLATSGD YRIFHVDKNG KRLSHIINPN NKRPISHNLA SISVVSDSAM
301 TADGLSTGLF VLGETEALRL AEQEKLAVFL IVRDKDGYRT AMSSEFAKLL
351 R*
```

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 401>: m111-1.seq

```
ATGCCGTCTG AAACACGCCT GCCGAACTTT ATCCGCGTCT TGATATTTGC
      CCTGGGTTTC ATCTTCCTGA ACGCCTGTTC GGAACAAACC GCGCAAACCG
  51
101
      TTACCCTGCA AGGCGAAACG ATGGGCACGA CCTATACCGT CAAATACCTT
      TCAAATAATC GGGACAAACT CCCCTCACCT GCCGAAATAC AAAAACGCAT
 151
      CGATGACGCG CTTAAAGAAG TCAACCGGCA GATGTCCACC TATCAGCCCG
      ACTCCGAAAT CAGCCGGTTC AACCAACACA CAGCCGGCAA GCCCCTCCGC
 251
301
      ATTTCAAGCG ACTTCGCACA CGTTACTGCC GAAGCCGTCC GCCTGAACCG
 351
      CCTGACACAC GGCGCGCTGG ACGTAACCGT CGGCCCCTTG GTCAACCTTT
 401
      GGGGATTCGG CCCCGACAAA TCCGTTACCC GTGAACCGTC GCCGGAACAA
      ATCAAACAGG CGGCATCTTA TACGGGCATA GACAAAATCA TTTTGAAACA
451
      AGGCAAAGAT TACGCTTCCT TGAGCAAAAC CCACCCCAAG GCCTATTTGG
501
551
      ATTTATCTTC GATTGCCAAA GGCTTCGGCG TTGATAAAGT TGCGGGCGAA
      CTGGAAAAAT ACGGCATTCA AAATTATCTG GTCGAAATCG GCGGCGAGTT
 601
 651
      GCACGGCAAA GGCAAAAACG CGCGGGGGA ACCGTGGCGC ATCGGTATCG
 701
      AGCAGCCCAA TATCGTCCAA GGCGGCAATA CGCAGATTAT CGTCCCGCTG
     AACAACCGTT CGCTTGCCAC TTCCGGCGAT TACCGTATTT TCCACGTCGA
TAAAAACGGC AAACGCCTCT CCCATATCAT CAACCCGAAC AACAAACGAC
 751
801
 851
      CCATCAGCCA CAACCTCGCC TCCATCAGCG TGGTCGCAGA CAGTGCGATG
901
      ACGGCGGACG GCTTGTCCAC AGGATTATTC GTATTGGGCG AAACCGAAGC
      CTTAAAGCTG GCAGAGCGCG AAAAACTCGC TGTTTTCCTG ATTGTCAGGG
951
      ATAAAGGCGG CTACCGCACC GCCATGTCTT CCGAATTTGA AAAACTGCTC
1001
1051
      CGCTAA
```

This corresponds to the amino acid sequence <SEQ ID 402; ORF 111-1>: m111-1.pep

```
1 MPSETRLPNF IRVLIFALGF IFLNACSEQT AQTVTLQGET MGTTYTVKYL
51 SNNRDKLPSP AEIQKRIDDA LKEVNRQMST YQPDSEISRF NQHTAGKPLR
101 ISSDFAHVTA EAVRLNRLTH GALDVTVGPL VNLWGFGPDK SVTREPSPEQ
151 IKQAASYTGI DKIILKQGKD YASLSKTHPK AYLDLSSIAK GFGVDKVAGE
201 LEKYGIQNYL VEIGGELHGK GKNARGEPWR IGIEQPNIVQ GGNTQIIVPL
251 NNRSLATSGD YRIFHVDKNG KRLSHIINPN NKRPISHNLA SISVVADSAM
301 TADGLSTGLF VLGETEALKL AEREKLAVFL IVRDKGGYRT AMSSEFEKLL
```

ml11-1/g111-1 96.6% identity in 351 aa overlap

BNSDOCID: <WO___9957280A2_|_>

	10	20	30	40	50	60	
	70	80	90	100	110	120	
ml11-1.pep	AEIQKRIDDALKEVNI	[]][]]		1111111111	1111111111	11111	
g111-1	AKIQKRIDDALKEVNI 70	80 80	90	100	110	120	
444 4	130	140	150	160	170	180	
ml11-1.pep	GALDVTVGPLVNLWG	[[]]	111111111	1111111111	1:11111111	11111	
g111-1	GALDVTVGPLVNLWG	140	150	160	170	180	
m111-1 pep	190 AYLDLSSIAKGFGVD	200 KVAGELEKYO	210	220	230 RGEPWRIGIE	240 OPNIVO	
mll1-1.pep		[ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ] [ ]		1111111111	:111111111	:	
g111-1	190	200	210	220	230	240	
m111-1.pep	250 GGNTQIIVPLNNRSL	260 ATSGDYRIFF	270 WDKNGKRLS	280 HIINPNNKRP	290	300 VADSAM	
g111-1			11111111	111111111	11111111111	1:1111	
giii-i	250	260	270	280	290	300	
m111-1.pep	310 TADGLSTGLFVLGET	320 EALKLAERER	330	340 KGGYRTAMSS	350 EFEKLLRX		
	TADGLETGLEVEGET	111:111:11	1111111111	1 11111111	11 11111		
g111-1	310	320	330	340	350		
hypothetical lipoprotein, Score = 34	JL HAEIN HYPOTHE	- Haemophi philus inf pect = 2e-	llus influ fluenzae F -95	enzae (str d] Length	ain Rd KW2 = 346	1074292 pir  C64144 0) >gi 1573128 (U3270 8 (1%)	2)
	NACSEQTAQTVTLOGE			SPAEIXKRID			
	, AC ++T + ++L G+ AACQKET-KVISLSGK						
	DSEISRFNQHT-AGKP			RLTHGALDVTV -+T GALDVTV			
	DSELSRFNQNTQVNTP						
Query: 142 V	TREPSPEQIKQAASYT		SKDYASLSKT	HPKAYLDLSS P+ Y+DLSS	IAKGFGVDKV	AGEL 201 A +I.	
Sbjct: 135 P	PEKOPTPEQLAERQAWV						
	CKYGIQNYLVEIGGELH	GKGKNARGEI KGKN G+1	PWRIGIEQPN PW+1 IE+P	IVQGGNTQII + +	VPLNNRSLAT	SGDY 261 SGDY	
	QLNAQNYMVEIGGEIR						
	RIFHVDKNGKRLSHIIN RI+ ++NGKR +H I+						
	RIY-FEENGKRFAHEID						
	REKLAVFLIVRDKGGY C+ LAV+LI+R G+						
Sbjct: 314 EKNNLAVYLIIRTDNGFVTKSSSAFKKL 341							
The following partial DNA sequence was identified in N. meningitidis <seq 403="" id="">: a111-1.seq</seq>							

1 ATGCCGTCTG AAACACGCCT GCCGAACTTT ATCCGCACCT TGATATTTGC
51 CCTGAGTTTT ATCTTCCTGA ACGCCTGTTC GGAACAAACC GCGCAAACCG
101 TTACCCTGCA AGGTGAAACG ATGGGCACGA CCTATACCGT CAAATACCTT
151 TCAAATAATC GGGACAAACT CCCCTCACCT GCCGAAATAC AAAAGCGCAT
201 CGATGACGCG CTTAAAAGAAG TCAACCGGCA GATGTCCACC TATCAGCCCG 251 ACTCCGAAAT CAGCCGGTTC AACCAACAC CAGCCGGCAA GCCCCTCCGC
301 ATTTCAAGCG ACTTCGCACA CGTTACTGCC GAAGCCGTCC ACCTGAACCG

35	51	CCTGACACAC	GGCGCGCTGG	ACGTAACCGT	CGGCCCCTTG	GTCAACCTTT
40	01	GGGGATTCGG	CCCCGACAAA	TCCGTTACCC	GTGAACCGTC	GCCGGAACAA
45	51	ATCAAACAAG	CAGCATCTTA	TACGGGCATA	GACAAAATCA	TTTTGAAACA
50	)1	AGGCAAAGAT	TACGCTTCCT	TGAGCAAAAC	CCACCCCAAG	GCCTATTTGG
55	51	ATTTATCTTC	GATTGCCAAA	GGCTTCGGCG	TTGATAAAGT	TGCGGGCGAA
60	)1	CTGGAAAAAT	ACGGCATTCA	AAATTATCTG	GTCGAAATCG	GCGGCGAGTT
65	51	GCACGGCAAA	GGCAAAAACG	CGCGCGGCGA	ACCTTGGCGC	ATCGGCATCG
70	)1	AACAGCCCAA	CATCGTCCAA	GGCGGCAATA	CGCAGATTAT	CGTCCCGCTG
75	51	AACAACCGTT	CGCTTGCCAC	TTCCGGCGAT	TACCGTATTT	TCCACGTCGA
80	)1	TAAAAGCGGC	AAACGCCTCT	CCCATATCAT	TAATCCGAAC	AACAAACGAC
85	51	CCATCAGCCA	CAACCTCGCC	TCCATCAGCG	TGGTCGCAGA	CAGTGCGATG
90	)1	ACGGCGGACG	GCTTGTCCAC	AGGATTATTC	GTATTGGGCG	AAACCGAAGC
95	51	CTTAAAGCTG	GCAGAGCGCG	AAAAACTCGC	TGTTTTCCTG	ATTGTCAGGG
100	)1	ATAAAGGCGG	CTACCGCACC	GCCATGTCTT	CCGAATTTGA	AAAACTGCTC
105	51	CGCTAA				

### This corresponds to the amino acid sequence <SEQ ID 404; ORF 111-1.a>: a111-1.pep

- 1 MPSETRLPNF IRTLIFALSF IFLNACSEQT AQTVTLQGET MGTTYTVKYL
  51 SNNRDKLPSP AEIQKRIDDA LKEVNRQMST YQPDSEISRF NQHTAGKPLR
  101 ISSDFAHVTA EAVHLNRLTH GALDVTVGPL VNLWGFGPDK SVTREPSPEQ

  - 151 IKQAASYTGI DKIILKQGKD YASLSKTHPK AYLDLSSIAK GFGVDKVAGE 201 LEKYGIQNYL VEIGGELHGK GKNARGEPWR IGIEQPNIVQ GGNTQIIVPL
  - 251 NNRSLATSGD YRIFHVDKSG KRLSHIINPN NKRPISHNLA SISVVADSAM
  - 301 TADGLSTGLF VLGETEALKL AEREKLAVFL IVRDKGGYRT AMSSEFEKLL
  - 351 R*

#### al11-1/ml11-1 98.9% identity in 351 aa overlap

	10	20	30	40	50	60
all1-1.pep	MPSETRLPNFIRTLIF					
m111-1	MPSETRLPNFIRVLIF	ALGFIFLNAC 20	SEQTAQTVTL 30	QGETMGTTYT 40	VKYLSNNRDK 50	LPSP 60
	10	20	30	40	30	60
	70	80	90	100	110	120
a111-1.pep	AEIQKRIDDALKEVNR					
m111-1	AEIQKRIDDALKEVNR 70	QMSTYQPDSE 80	ISRFNOHTAG 90	KPLRISSDFA 100		
	70	80	90	100	110	120
	130	140	150	160	170	180
al11-1.pep	GALDVTVGPLVNLWGF	GPDKSVTREP	SPEQIKQAAS	YTGIDKIILK	QGKDYASLSK	THPK
	413111111111111111111111111111111111111					
m111-1	GALDVTVGPLVNLWGF				_	
	130	140	150	160	170	180
	190	200	210	220	230	240
a111-1.pep	AYLDLSSIAKGFGVDK					
• •		1111111111	111111111	1111111111	ниний	HH
m111-1	AYLDLSSIAKGFGVDK		_		_	_
	190	200	210	220	230	240
	250	260	270	280	290	300
a111-1.pep	GGNTOIIVPLNNRSLA					
	11111111111111111	111111111	11:111111	1111111111	1111111111	1111
m111-1	GGNTQIIVPLNNRSLA					DSAM
	250	260	270	280	290	300
	310	320	330	340	350	
all1-1.pep	TADGLSTGLFVLGETE					
1.рер	111111111111111111					
m111-1	TADGLSTGLFVLGETE					
	310	320	330	340	350	

### The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 405>: gl14.seq

- ATGGCTTCCA TCACTTCGCC GCTGCACGGG GCGCAGCAGG AATGCAGCAA 1
- 51 GACTTTTTA TGTCCGCCGG GCGGGACGAG TATGGGGCGG TCAATGTCGG

حرز BNSDOCID: <WO___9957280A2_j

a • • • ...

WO 99/57280

```
101 TAACGGTAGG TTTGTTTTGT GTTTCCATTA ACTTAACAAT ATCTGTCGAA
          151 TACGGTCAAA GCGGCTATTT TACCAGAGCC GCCGAATGTA AAACAGGGTG
               TCAGGGCATC AGCCCGAGCT GCCTGAACGA ACGGACGGTT TGCGAGGTAA
          201
               CGATAAAATG GTCGAGCAGC GAAACATCAA CCAGCGACAT GGCCTGTGCC
          251
          301 AGCCGCCTTG TGAACATGAT GTCTTCCTGC GAAGGTTCAG GCGAGCCGCC
          351 CGGATGGTTG TGCGCGATAA TCAGGCTGTC GGCATATTCG TCCAATGCCA
               GTTTGACGAT TTCGCGGATG TAA
This corresponds to the amino acid sequence <SEQ ID 406; ORF 114.ng>:
     gl14.pep
              MASITSPLHG AQQECSKTFL CPPGGTSMGR SMSVTVGLFC VSINLTISVE
            1
              YGQSGYFTRA AECKTGCQGI SPSCLNERTV CEVTIKWSSS ETSTSDMACA
           51
              SRLVNMMSSC EGSGEPPGWL CAIIRLSAYS SNASLTISRM *
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 407>:
     mll4.seq
              ATGGCTTCCA TCACTTCGCC GCTGCACGGG GCGCACAGAG AATGCAGCAA
              GACTTTTTTA TGTCCACCGG GCGGGACGAG TATAGGGCGG TCAATGTCGG
           51
         101 TAACGGTAGG TTTGTTTTGT GTTTCCATTA ACTTAACAAT ATCTGTTGAA
         151 TACGGTTGAA GCGGCTATTT TATCAGAGCC GCCGCATGTA AAACAGAGTG
         201 TCAGGGCATC AACCCGAGCT GTCTGAACGA ACAGACGCTT TGCGAKGTAA
         251 CGATAAAATG GTCGAGCAGC GACACATCGA CCAGCGACAT TGCCTGTGCC
         301 AGCCGCCTTG TGAACATGAT GTCTTCCTGC GAArGTTCsG GCGAGCCGcC
         351 CGGATGGTTG TGCGCAATAA TCAGGCTGTC GGCATATTCG TCCAATGCCA
         401 GTTTGACGAT TTCGCGGATG TAA
This corresponds to the amino acid sequence <SEQ ID 408; ORF 114>:
    m114.pep
              MASITSPLHG AHRECSKTFL CPPGGTSIGR SMSVTVGLFC VSINLTISVE
           1
              YGXSGYFIRA AACKTECQGI NPSCLNEQTL CXVTIKWSSS DTSTSDIACA
          51
         101 SRLVNMMSSC EXSGEPPGWL CAIIRLSAYS SNASLTISRM *
    ml14/g114 90.0% identity over a 140 aa overlap
                        10
                                  20
                                           30
                                                     40
                 MASITSPLHGAHRECSKTFLCPPGGTSIGRSMSVTVGLFCVSINLTISVEYGXSGYFIRA
    ml14.pep
                 MASITSPLHGAQQECSKTFLCPPGGTSMGRSMSVTVGLFCVSINLTISVEYGQSGYFTRA
    g114
                                  20
                                           30
                                                     40
                                                              50
                                                                        60
                        70
                                  80
                                           90
                                                    100
                                                             110
                AACKTECQGINPSCLNEQTLCXVTIKWSSSDTSTSDIACASRLVNMMSSCEXSGEPPGWL
    m114.pep
                 AECKTGCQGISPSCLNERTVCEVTIKWSSSETSTSDMACASRLVNMMSSCEGSGEPPGWL
    g114
                        70
                                 80
                                           90
                                                   100
                                                             110
                       130
                                 140
    ml14.pep
                CAIIRLSAYSSNASLTISRMX
                111111111111111111111111
   g114
                CAIIRLSAYSSNASLTISRMX
                       130
                                140
```

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 409>: all4.seq

1	ATGCCGGAGG	CAAGCATCGC	CTCCATCACT	TOGOCGCTGC	ACGGGGCGCA
51	ACAGGAATGC	AGCAAGACTT	ጥጥጥጥልጥርጥርር	5000000100	ACGGGGGGCA
101	GGCGGTCAAT	GTCGGTAACC	CHACCERE		ACGAGTATGG
151	A C C A M A M C M C	GTCGGTAACG	GIAGGITTGT	TTTGTGTTTC	CATTAACTTA
	ACGATATCTG	TCGAATACGG	TTGAAGCGGC	TATTTTATCA	GAGCCGCCGC
201	AIGIAAAACA	GGGTGTCAGG	GCATCAGCCC	GAGCTGCCTC	776677667
251	CGGTTTGCGC	CGTTACGATA	AAATGGTCCA	GCACCCACAC	ADDOMADOR
301	GACATTGCCT	GTGCCAGCCG	CCTTCTCTTT	ATTENTON	ATCGACCAGC
351	TTCGGGCGAG	CCCCCCCAM	CCTTGTGAAC	ATGATGTCTT	CCTGCGAAGG
401	TTCGGGCGAG	CCGCCCGGAT	GGTTGTGCGC	GATAATCAGG	CTGTCGGCAT
401	ATTCGTCCAA	TGCCAGTTTG	ACAATTTCAC	GGATGTAA	

This corresponds to the amino acid sequence <SEQ ID 410; ORF 114.a>: a114.pep

1 MPEASIASIT SPLHGAQQEC SKTFLCPPGG TSMGRSMSVT VGLFCVSINL
51 TISVEYG*SG YFIRAAACKT GCQGISPSCL NERTVCAVTI KWSSSDTSTS
101 DIACASRLVN MMSSCEGSGE PPGWLCAIIR LSAYSSNASL TISRM*

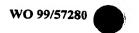
### m114/a114 92.9% identity in 140 aa overlap

		10	20	30	40	50
mll4.pep	MASIT	SPLHGAHRE	CSKTFLCPPGG	TSIGRSMSV	TVGLFCVSIN	LTISVEYGXSG
	:1111	111111::11		11:11111	111111111	
a114	MPEASIASIT	SPLHGAQQE			TVGLFCVSIN	LTISVEYGXSG
	10	20	30	) 4	0 5	0 60
	60	7.0	0.0	0.0	100	110
	60	70	80	90	100	110
m114.pep	YFIRAAACKT	ECQGINPSCI	LNEQTLCXVTI	KWSSSDTST	SDIACASRLV	NMMSSCEXSGE
	111111111	1111:111	: :		111111111	1111111
a114	YFIRAAACKT	GCQGISPSCI	LNERTVCAVTI	KWSSSDTST	SDIACASRLV	NMMSSCEGSGE
	70	80	90	10	0 11	0 120
	120	130	140			
m114.pep	PPGWLCAIIR					
mara.pep	1111111111		1111111			
	DDCWL CATED	TONUCCHACI	TOTODAY			
a114	PPGWLCAIIR					
	130	140	)			

### The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 411>:

```
g117.seq
         atggtcgacg aactcgacCT GCTGCCCGAT GCCGTCGCCG CCACCCTGCT
      1
         TGCCGACATC GGACGCTACG TCCCCGATTG GAACCTATTG GTTTCCGAGC
     51
    101 GCTGCAACAG CACCGTCGCC GAGCTGGTCA AAGGTGtgga CGAAGTGCAG
    151 AAACTTACCC ACTTCGCCCG GGTGGACAGC CTCGCCACGC CGGAAGAACG
    201 CGCACAGCAA GCGGAAACCA TGCGGAAAAT GCTGCTGGCg atggttaccg
    251 Acatecgegt egtaTTAATC AAACTGGCGA TGCGTacgeg caceCTGcta
    301 ttTTtaaGCA ACGCCCCCGA CAGCCCTGAA AAACgcgccG TCgccaaAga
    351 aacceTCGAC ATCTTCGCCC CGCTCGCCAA CCGCTTGGGC GTGTGGCAGC
    401 TCAAATGGCA GCTCGAAGAT TTGGGCTTCC GCCATCAAGA ACCCGAAAAA
         TACCGCGAAA TCGCCCTGCT TTTGGACGAA AAACGCACCG AACGCCTCGA
    451
         ATACATCGAA AACTTCCTCG ATATCCTGCG TACGGAACTC AAAAAATACA
         ATATCCACTT TGAAGTCGCC GGCCGTCCGA AACACATCTA CTCCATTTAC
    601 AAAAAATGG TGAAGAAAAA ACTCAGCTTC GACGgccTGT TCGACATCCG
    651 CGCCGTGCGG ATTCTGGTCG ATACCGTCCC CGaGTGTTAC ACCACGCTGG
    701 gcaTCGTCCA CAGCCTCTGG CAGCCCATTC CCGGCGagtt CGAcgactAC
    751 ATCGCCAACC CCAAAGqcaA CGqttATAAA AGtTTGCACA CCGTCATCGT
    801 cqqccGGAa qacaaaggtg tggaaGtgCA AATCCGCACC TTCGAtatGC
    851 accAATTCaa CgaatTcggT gtcgccgCCC ACTGGCGtta caaagaaggc
    901 ggcaaaggcg attccGCCtA cgaacaaAAA ATcgccTggt TGCgccaACT
    951 CTTGGACTGG CGCGAAAATA TGGCGGAAAG CGGCAAGGAA GACCTCGCCG
   1001 CCGCCTTCAA AACCGAGCTT TTCAACGACA CGATTTATGT TTTGACCCCG
   1051 CACGGCAAAG TCCTCTCTT GCCAACGGGC GCAACCCCCA TCGACTTCGC
   1101
         CTACGCCTG CACAGCAGCA TeggCGACCG CTGCCGGGGC GCGAAAGTCG
   1151 AaggGCAGAT TGTGCCGCTG TCCACCCCGC TCGAAAACGG ACAGCGCGTC
         GAAATcatta cCGCcaaAGA AGGGCATCCT TCCGTCAACT GGCTTTACGA
   1251 AGGCtgGGtc aAATCCGGCA AGGCCATCGG caaAATCCGC GCCTAcatCC
   1301 GCCAGcaaAa cgCcgaCACC GTGCGCGAAG AAGGCCGTGT CCAACTCGAC
   1351 AAGCAGCTTG CCAAACTCAC GCCCAAACCC AACCTGCAAG AGCTTgccga
   1401 aaATCTCGGC tacaaAAAGC cagaagacct ctacacCGCc gtcggacaaq
   1451 gcgaaatttc caaccgcgcc atCcaaaaag cctgcggcac GCTgaacgaa
   1501 CCGCCCCCG TGCCCGTCAG CGCAACCACC ATCGTCAAAC AGTCCAAAAT
```

BNSDOCID: <WO___9957280A2_J_>



```
1551 CAAAAAAGGT GGCAAAACCG GCGTGCTCAT CGACGGCGAA GACGGCTTGA
1601 TGACCACGCT TGCCAAATGC TGCAAACCCG CGCCGCCCGA CGATATTGCC
1651 GGCTTCGTTA CCCGCGAGCG CGCATTTCC GTCCACCGCA AAACCTGCCC
1701 CTCTTTCCGA CACCTTGCCG AACACGCGCC CGAAAAAGTA CTGGACGCAA
1751 GTTGGGCGGC GTTGCAGGAA GGGCAAGTGT TCGCCGTCGA TATCGAAATC
1801 CGCGCCCAAG ACCGCTCCGG GCTTTTGCGC GACGTATCCG ACGCGCTCGC
1851 CCGCCACAAA CTCAACGTTA CCGCCGTGCA AACCCAGTCC CGCGACTTGG
1901 AAGCCAGCAT GAGGTTCACG CTCGAAGTCA AACAAGCCAA CGACCTCCCG
1951 CGCGTCCTCG CCGCCTCGG CGATGTCAAA GGCGTATTGA GCGTTACCCG
```

## This corresponds to the amino acid sequence <SEQ ID 412; ORF 117.ng>:

```
MVDELDLLPD AVAATLLADI GRYVPDWNLL VSERCNSTVA ELVKGVDEVQ

1 KLTHFARVDS LATPEERAQQ AETMRKMLLA MVTDIRVVLI KLAMRTRTLL

101 FLSNAPDSPE KRAVAKETLD IFAPLANRLG VWQLKWQLED LGFRHQEPEK

151 YREIALLLDE KRTERLEYIE NFLDILRTEL KKYNIHFEVA GRPKHIYSIY

201 KKMVKKKLSF DGLFDIRAVR ILVDTVPECY TTLGIVHSLW QPIPGEFDDY

251 IANPKGNGYK SLHTVIVGPE DKGVEVQIRT FDMHQFNEFG VAAHWRYKEG

301 GKGDSAYEQK IAWLRQLLDW RENMAESGKE DLAAAFKTEL FNDTIYVLTP

351 HGKVLSLPTG ATPIDFAYAL HSSIGDRCRG AKVEGQIVPL STPLENGQRV

401 EIITAKEGHP SVNWLYEGWV KSGKAIGKIR AYIRQQNADT VREEGRVQLD

401 KQLAKLTPKP NLQELAENLG YKKPEDLYTA VGQGEISNRA IQKACGTLNE

501 PPPVPVSATT IVKQSKIKKG GKTGVLIDGE DGLMTTLAKC CKPAPPDDIA

501 FOVTRERGIS VHRKTCPSFR HLAEHAPEKV LDASWAALQE GQVFAVDIEI

601 RAQDRSGLLR DVSDALARHK LNVTAVQTQS RDLEASMRFT LEVKQVNDLP
```

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 413>: m117.seq (partial)

```
..GTGAAACTCA AGAAATACAA TGTCCATTTC GAAGTCGCCG GCCGCCCGAA
   1
        ACACATCTAC TCCATTTACA AAAAAATGGT GAAGAAAAAA CTCAGCTTCG
        ACGGCCTCTT TGACATCCGC GCCGTGCGAA TTCTGGTTGA TACCGTCCCC
        GAGTGTTACA CCACGCTGGG TATCGTCCAC AGCCTCTGGC AGCCCATTCC
 151
 201
        CGGCGAGTTC GACGACTACA TCGCCAATCC CAAAGGCAAC GGCTATAAAA
 251
        GTTTGCACAC CGTCATCGTC GGCCCGGAAG ACAAAGGCGT GGAAGTACAA
        ATCCGCACCT TCGATATGCA CCAATTCAAC GAATTCGGTG TCGCCGCCCA
 301
        CTGGCGTTAC AAAGAGGGCG GCAAGGGCGA TTCCGCCTAC GAACAGAAAA
 351
        TCGCCTGGTT GCGCCAACTC TTGGACTGGC GCGAAAACAT GGCGGAAAGC
 401
        GGCAAGGAAG ACCTCGCCGC CGCCTTCAAA ACCGAGCTTT TCAACGACAC
 451
 501
        GATTTATGTT TTGACCCCGC ACGGCAAAGT CCTCTCCCTG CCCACGGGCG
        CGACCCCCAT CGACTTCGCC TACGCCCTGC ACAGCAGCAT CGGCGACCGT
 551
        TGCCGCGGTG CGAAAGTCGA AGGGCAGATT GTGCCGCTGT CCACCCCGCT
 601
        CGAAAACGGA CAGCGCGTCG AAATCATTAC CGCCAAAGAA GGGCATCCTT
 651
        CCGTCAACTG GCTTTACGAA GGCTGGGTCA AATCCAACAA GGCAATCGGC
 701
        AAAATCCGCG CCTACATCCG CCAGCAAAAC GCCGACACCG TGCGCGAAGA
 751
        AGGCCGCGTC CAACTCGACA AACAGCTTGC CAAACTCACG CCCAAACCCA
 801
        ACCTGCAAGA GCTTGCCGAA AATCTCGGCT ACAAAAAGCC AGAAGACCTC
 851
        TACACCGCCG TCGGACAAGG CGAAATTTCC AACCGCGCCA TCCAAAAAGC
 901
        CTGCGGCACG CTGAACGAAC CGCCGCCCGT ACCCGTCAGC GAAACCACCA
 951
        TCGTCAAACA GTCCAAAATC AAAAAAGGCG GCAAAAACGG CGTGCTCATC
1001
        GACGGCGAAG ACGGTCTGAT GACCACGCTT GCCAAATGCT GCAAACCCGC
1051
        GCCGCCCGAC GATATTATCG GCTTCGTTAC CCGCGAGCGC GgCATTTCAG
1101
        TGCACCGCAA AWYYTKCYCG TCTTTCCAAC ACCTCGCCGA ACACGCGCCC
1151
       GAWAAAGTGC TGGACGCAAG CTGGGCGGCA TTGCAGGAAG GACAAGTATT
1201
       CGCCGTCGAT ATCGAAATCC GCGCCCAAGA CCGCTCCGGG CTTTTGCGCG
1251
       ACGTATCCGA CGCGCTCGCC CGCCACAAAC TCAACGTTAC CGCCGTGCAA
1301
       ACCCAGTCCC GCGACTTGGA AGCCAGCATG AGGTTCACGC TCGAAGTCAA
1351
       ACAAGTCAAC GACCTCCCGC GCGTCCTCGC CAGCCTCGGC GACGTCAAAG
1401
       GCGTATTGAG CGTTACCCGG CTTTAA
1451
```

This corresponds to the amino acid sequence <SEQ ID 414; ORF 117>:

1	VKLKKYNVHF	EVAGRPKHIY	SIYKKMVKKK	LSFDGLFDIR	AVRILVDTVP
51	ECYTTLGIVH	SLWQPIPGEF	DDYIANPKGN	GYKSLHTVIV	GPEDKGVEVQ
101	IRTFDMHQFN	EFGVAAHWRY	KEGGKGDSAY	EQKIAWLRQL	LDWRENMAES
151	GKEDLAAAFK	TELFNDTIYV	LTPHGKVLSL	PTGATPIDFA	YALHSSIGDR
201	CRGAKVEGQI	VPLSTPLENG	QRVEIITAKE	GHPSVNWLYE	GWVKSNKAIG
251	KIRAYIRQQN	ADTVREEGRV	QLDKQLAKLT	PKPNLQELAE	NLGYKKPEDL
301	YTAVGQGEIS	NRAIQKACGT	LNEPPPVPVS	ETTIVKQSKI	KKGGKNGVLI
351	${\tt DGEDGLMTTL}$	AKCCKPAPPD	DIIGFVTRER	GISVHRKXXX	SFQHLAEHAP
401	XKVLDASWAA	LQEGQVFAVD	IEIRAQDRSG	LLRDVSDALA	RHKLNVTAVQ
451	TOSRDLEASM	RFTLEVKQVN	DLPRVLASLG	DVKGVLSVTR	L*

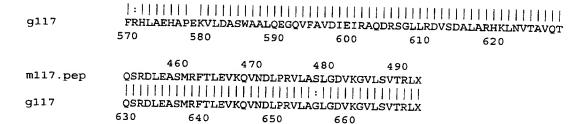
451 TOSRDLEASM RFTLEVKQVN DLPRVLASLG DVKGVLSVTR L*
Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae
ORF 117 shows 97.6% identity over a 490 aa overlap with a predicted ORF (ORF 117.ng) from *N. gonorrhoeae:* m117/g117

				10	20	30
m117.pep			VKI	LKKYNVHFEVAC	GRPKHIYSIYKK	MVKKKL
			:	:		111111
g117	EKYREIALL	LDEKRTERLEY	IENFLDILRTEI	LKKYNIHFEVAC	GRPKHIYSIYKK	MVKKKL
_	150	160 1	70 180	190	200	
	40	50	60	70	80	90
m117.pep	SFDGLFDIR	AVRILVDTVPE			ANPKGNGYKSL	
g117	SEDGLEDIR	AVRTIVOTVPE	CYTTLGTVHSL	MODI PGEFODY I	IANPKGNGYKSL	HULLIL
9			30 240	-	260	mitvivg
	210	220 2	240	250	200	
	100	110	120	130	140	150
-117					IAWLRQLLDWRE	
ml17.pep						
g117					AWLRQLLDWRE	NMAESG
	270	280 2	90 300	310	320	
	160	170	180	190	200	210.7
m117.pep	KEDLAAAFK				HSSIGDRCRGAK	
g117	KEDLAAAFK	TELFNDTIYVL	TPHGKVLSLPTO	GATPIDFAYAL	<b>ISSIGDRCRGA</b> K	VEGQIV
	330	340 3	50 360	370	380	
	220	230	240	250	260	270
m117.pep	PLSTPLENG	<b>ORVEIITAKEG</b>	HPSVNWLYEGW\	/KSNKAIGKIRA	YIRQQNADTVR	EEGRVO
g117	PLSTPLENG	ORVEIITAKEG	HPSVNWLYEGW	KSGKAIGKIR	YIRQQNADTVR	EEGRVO
_			10 420		440	
	280	290	300	310	320	330
m117.pep					QKACGTLNEPP	
g117						
gii,			70 480		500	PVPVSA
	450	400 4	/0 480	490	500	
	340	350	260	270	200	
			360	370	380	390
m117.pep					FVTRERGISVH	
					[	
g117					FVTRERGISVH	RKTCPS
	510	<b>52</b> 0 5	30 540	550	560	
	400	410	420	430	440	450
m117.pep	FQHLAEHAP:	XKVLDASWAAL	QEGQVFAVDIE:	RAQDRSGLLRI	OVSDALARHKLN	VTAVQT

BNSDOCID: <WO__9957280A2_l_>

WO 99/57280



The following partial DNA sequence was identified in N. meningitidis <SEQ ID 415>: a117.seq

```
ATGGTTCATG AACTCGACCT GCTCCCCGAT GCCGTCGCCG CCACCCTGCT
   51
       TGCCGACATC GGACGCTACG TCCCCGACTG GAACCTATTG GTTTCCGAAC
       GCTGCAACAG TACCGTCGCC GAGCTGGTCA AAGGTGTGGA CGAAGTGCAG
  101
  151
       AAACTCACCC ACTTCGCCCG GGTGGACAGC CTCGCCACGC CGGAAGAACG
  201
       CGCCCAGCAG GCAGAAACTA TGCGGAAAAT GCTGCTGGCG ATGGTTACCG
  251
       ACATCCGCGT CGTGTTAATC AAACTGGCGA TGCGTACGCG CACCCTGCAA
       TTTTTAAGCA ACGCCCCGA CAGCCCCGAA AAACGCGCCG TCGCCAAAGA
  301
       AACCCTCGAC ATCTTCGCCC CGCTCGCCAA CCGTTTGGGC GTGTGGCAGC
  351
       TCAAATGGCA GCTCGAAGAT TTGGGCTTCC GCCATCAAGA ACCCGAAAAA
 401
       TACCGCGAAA TCGCCCTGCT TTTGGACGAA AAACGCACCG AACGCCTCGA
 451
      ATACATCGAA AACTTCCTTA ATATCCTGCG TACGGAACTC AAAAAATACA
 501
 551
      ATATCCACTT TGAAGTCGCC GGCCGTCCGA AACACATCTA CTCCATTTAC
      AAAAAAATGG TGAAGAAAAA ACTCAGCTTC GACGGGTTGT TCGACATCCG
 601
      CGCCGTGCGG ATTCTGGTTG ATACCGTCCC CGAGTGTTAC ACCACACTGG
 651
 701
      GCATTGTCCA CAGCCTCTGG CAGCCCATTC CCGGCGAGTT CGACGACTAC
      ATCGCCAACC CGAAAGGCAA CGGCTATAAA AGTTTGCACA CCGTCATCGT
 751
 801
      CGGCCCGGAA GACAAAGGCG TGGAAGTGCA AATCCGCACC TTCGATATGC
      ACCAATTCAA CGAATTCGGT GTCGCCGCGC ACTGGCGTTA CAAAGAGGGC
 851
 901
      GGCAAAGGCG ATTCCGCCTA CGAACAAAAA ATCGCCTGGT TACGCCAACT
 951
      TTTGGACTGG CGCGAAAACA TGGCGGAAAG CGGCAAGGAA GACCTCGCCG
1001
      CCGCCTTCAA AACCGAGCTT TTCAACGACA CGATTTATGT TTTGACCCCG
      CACGGCAAAG TCCTCTCCCT GCCCACAGGC GCGACCCCCA TCGACTTCGC
1051
      CTACGCCCTG CACAGCAGCA TCGGCGACCG TTGCCGCGGT GCGAAAGTCG
1101
      AAGGGCAGAT TGTGCCGCTG TCCACCCCGC TCGAAAACGG ACAGCGTGTC
1151
      GAAATCATTA CCGCCAAAGA AGGGCATCCT TCCGTCAACT GGCTTTACGA
1201
      AGGCTGGGTC AAATCCAACA AGGCAATCGG CAAAATCCGC GCCTACATCC
1251
1301
      GCCAGCAAAA CGCCGACACC GTGCGCGAAG AAGGCCGCGT CCAACTCGAC
      AAACAGCTTG CCAAACTCAC GCCCAAACCC AACCTGCAAG AGCTTGCCGA
1351
     AAATCTCGGC TACAAAAAGC CAGAAGACCT CTACACCGCC GTCGGACAAG
1401
      GCGAAATTTC CAACCGCGCC ATCCAAAAAG CCTGCGGCAC GCTGAACGAA
1451
1501
      CCGCCGCCCG TACCCGTCAG CGAAACCACC ATCGTCAAAAC AGTCCAAAAT
     CAAAAAAGGC GGCAAAAACG GCGTGCTCAT CGACGGCGAA GACGGTCTGA
1551
      TGACCACGCT TGCCAAATGC TGCAAACCCG CGCCGCCCGA CGACATTGTC
1601
1651
      GGCTTCGTTA CCCGCGATCG CGGCATTTCG GTACACCGCA AAACCTGCCC
      CTCTTTCCGA CACCTCGCCG AACACGCGCC CGAAAAAGTA CTGGACGCAA
1701
      GTTGGGCGGC GTTGCAGGAA GGACAAGTGT TCGCCGTCGA TATCGAAATC
1751
      CGCGCCCAAG ACCGCTCCGG GCTTTTGCGC GACGTATCCG ACGCGCTCGC
1801
      CCGCCACAAA CTCAACGTTA CCGCCGTGCA AACCCAGTCC CGCGACTTGG
1851
      AAGCCAGCAT GAGGTTCACG CTCGAAGTCA AACAAGTTAC CGACCTCCCA
1901
      CGCGTCCTCG CCAGCCTCGG CGACGTCAAA GGCGTATTGA GCGTTACCCG
1951
2001
      GCTTTAA
```

This corresponds to the amino acid sequence <SEQ ID 416; ORF 117.a>: a117.pep

EVQ
ת דית
שישום
O T W
אמר
750
(II)
DRV
PE SI OD KE

401 451 501 551 601 651	EIITAKEGHP SVNWLYEGWV KSNKAIGKIR AYIRQQNADT VREEGRVQLD KQLAKLTPKP NLQELAENLG YKKPEDLYTA VGQGEISNRA IQKACGTLNE PPPVPVSETT IVKQSKIKKG GKNGVLIDGE DGLMTTLAKC CKPAPPDDIV GFVTRDRGIS VHRKTCPSFR HLAEHAPEKV LDASWAALQE GQVFAVDIEI RAQDRSGLLR DVSDALARHK LNVTAVQTQS RDLEASMRFT LEVKQVTDLP RVLASLGDVK GVLSVTRL*
m117/a117	98.0% identity in 490 aa overlap
m117.pep	10 20 30  VKLKKYNVHFEVAGRPKHIYSIYKKMVKKKL :    :
	150 160 170 180 190 200 40 50 60 70 80 90
m117.pep	SFDGLFDIRAVRILVDTVPECYTTLGIVHSLWQPIPGEFDDYIANPKGNGYKSLHTVIVG
a117	SFDGLFDIRAVRILVDTVPECYTTLGIVHSLWQPIPGEFDDYIANPKGNGYKSLHTVIVG 210 220 230 240 250 260
ml17.pep	100 110 120 130 140 150 PEDKGVEVQIRTFDMHQFNEFGVAAHWRYKEGGKGDSAYEQKIAWLRQLLDWRENMAESG
a117	PEDKGVEVQIRTFDMHQFNEFGVAAHWRYKEGGKGDSAYEQKIAWLRQLLDWRENMAESG 270 280 290 300 310 320
m117.pep	160 170 180 190 200 210 KEDLAAAFKTELFNDTIYVLTPHGKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIV
a117	KEDLAAAFKTELFNDTIYVLTPHGKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIV 330 340 350 360 370 380
m117.pep	220 230 240 250 260 270 PLSTPLENGQRVEIITAKEGHPSVNWLYEGWVKSNKAIGKIRAYIRQQNADTVREEGRVQ
a117	PLSTPLENGQRVEIITAKEGHPSVNWLYEGWVKSNKAIGKIRAYIRQQNADTVREEGRVQ 390 400 410 420 430 440
m117.pep	280 290 300 310 320 330 LDKQLAKLTPKPNLQELAENLGYKKPEDLYTAVGQGEISNRAIQKACGTLNEPPPVPVSE
a117	LDKQLAKLTPKPNLQELAENLGYKKPEDLYTAVGQGEISNRAIQKACGTLNEPPPVPVSE 450 460 470 480 490 500
m117.pep	340 350 360 370 380 390 TTIVKQSKIKKGGKNGVLIDGEDGLMTTLAKCCKPAPPDDIIGFVTRERGISVHRKXXXS
a117	TTIVKQSKIKKGGKNGVLIDGEDGLMTTLAKCCKPAPPDDIVGFVTRDRGISVHRKTCPS 510 520 530 540 550 560
m117.pep	400 410 420 430 440 450 FQHLAEHAPXKVLDASWAALQEGQVFAVDIEIRAQDRSGLLRDVSDALARHKLNVTAVQT
a117	:
m117.pep	460 470 480 490 QSRDLEASMRFTLEVKQVNDLPRVLASLGDVKGVLSVTRLX
a117	QSRDLEASMRFTLEVKQVTDLPRVLASLGDVKGVLSVTRLX 630 640 650 660

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 417>:

```
ATGACCGCCA TCAGCCCGAT TCAAGACACG CAAAGCGCGA CCCTGCAAGA
   51 ATTGCGCGAA TGGTTCGACA GCTACTGCGC CGCTCTGCCG GACAACGATA
  101 AAAACCTCAT CGGTACCGCA TGGTCGCTGG CGCAGGAACA TTATCCTGCC
      GATGCCGCCA CGCCGTATGG CGAGCCGCTG CCCGACCACT TCCTCGGCGC
  151
  201 GGCGCAAATG GTCGACGAAC TCGACCTGCT GCCCGATGCC GTCGCCGCCA
       CCCTGCTTGC CGACATCGGA CGCTACGTCC CCGATTGGAA CCTATTGGTT
  251
  301
       TCCGAGCGCT GCAACAGCAC CGTCGCCGAG CTGGTCAAAG GTGTGGACGA
  351
      AGTGCAGAAA CTTACCCACT TCGCCCGGGT GGACAGCCTC GCCACGCCGG
  401 AAGAACGCGC ACAGCAAGCG GAAACCATGC GGAAAATGCT GCTGGCGATG
      GTTACCGACA TCCGCGTCGT ATTAATCAAA CTGGCGATGC GTACGCGCAC
  451
 501 CCTGCAATTT TTAAGCAACG CCCCGACAG CCCTGAAAAA CGCGCCGTCG
551 CCAAAGAAAC CCTCGACATC TTCGCCCCGC TCGCCAACCG CTTGGGCGTG
  601 TGGCAGCTCA AATGGCAGCT CGAAGATTTG GGCTTCCGCC ATCAAGAACC
  651 CGAAAAATAC CGCGAAATCG CCCTGCTTTT GGACGAAAAA CGCACCGAAC
      GCCTCGAATA CATCGAAAAC TTCCTCGATA TCCTGCGTAC GGAACTCAAA
  701
  751
      AAATACAATA TCCACTTTGA AGTCGCCGGC CGTCCGAAAC ACATCTACTC
 801
      CATTTACAAA AAAATGGTGA AGAAAAAACT CAGCTTCGAC GGCCTGTTCG
      ACATCCGCGC CGTGCGGATT CTGGTCGATA CCGTCCCCGA GTGTTACACC
 851
      ACGCTGGGCA TCGTCCACAG CCTCTGGCAG CCCATTCCCG GCGagttCGA
 901
 951
      CGACTACATC GCCAACCCCA AAGGCAACGG ttATAAAAGt TTGCACACCG
1001
      TCATCGTcgg cccGGAagaa aaaggtgtgg aagtgcAAAT CCGCACCTTC
      GATATGCACC AATTCAACGA ATTCGGTGTC GCCGCCCACT GGCGTTACAA
1051
1101
      AGAAGGCGGC AAAGGCGATT CCGCCTACGA ACAAAAAATC GCCTGGTTGC
      GCCAACTCTT GGACTGGCGC GAAAATATGG CGGAAAGCGG CAAGGAAGAC
1151
      CTCGCCGCCG CCTTCAAAAC CGAGCTTTTC AACGACACGA TTTATGTTTT
1201
      GACCCCGCAC GGCAAAGTCC TCTCTCTGCC AACGGGCGCA ACCCCCATCG
1251
1301 ACTTCGCCTA CGCCCTGCAC AGCAGCATCG GCGACCGCTG CCGGGGCGCG
      AAAGTCGAAG GGCAGATTGT GCCGCTGTCC ACCCCGCTCG AAAACGGACA
1351
      GCGCGTCGAA ATCATTACCG CCAAAGAAGG GCATCCTTCC GTCAACTGGC
1401
      TTTACGAAGG CTGGGTCAAA TCCGGCAAGG CCATCGGCAA AATCCGCGCC
1451
      TACATCCGCC AGCAAAACGC CGACACCGTG CGCGAAGAAG GCCGTGTCCA
1501
1551 ACTCGACAAG CAGCTTGCCA AACTCACGCC CAAACCCAAC CTGCAAGAGC
     TTGCCGAAAA TCTCGGCTAC AAAAAGCCAG AAGACCTCTA CACCGCCGTC
1601
1651 GGACAAGGCG AAATTTCCAA CCGCGCCATC CAAAAAGCCT GCGGCACGCT
     GAACGAACCG CCGCCCGTGC CCGTCAGCGC AACCACCATC GTCAAACAGT
1701
1751
      CCAAAATCAA AAAAGGTGGC AAAACCGGCG TGCTCATCGA CGGCGAAGAC
1801 GGCTTGATGA CCACGCTTGC CAAATGCTGC AAACCCGCGC CGCCCGACGA
     TATTGCCGGC TTCGTTACCC GCGAGCGCGG CATTTCCGTC CACCGCAAAA
1851
     CCTGCCCCTC TTTCCGACAC CTTGCCGAAC ACGCGCCCGA AAAAGTACTG
1901
1951
     GACGCAAGTT GGGCGGCGTT GCAGGAAGGG CAAGTGTTCG CCGTCGATAT
2001
     CGAAATCCGC GCCCAAGACC GCTCCGGGCT TTTGCGCGAC GTATCCGACG
     CGCTCGCCCG CCACAAACTC AACGTTACCG CCGTGCAAAC CCAGTCCCGC
     GACTTGGAAG CCAGCATGAG GTTCACGCTC GAAGTCAAAC AAGTCAACGA
     CCTCCCGCGC GTCCTCGCCG GCCTCGGCGA TGTCAAAGGC GTATTGAGCG
     TTACCCGGCT TTAA
```

This corresponds to the amino acid sequence <SEQ ID 418; ORF 117-1.ng>:

```
1 MTAISPIQDT QSATLQELRE WFDSYCAALP DNDKNLIGTA WSLAQEHYPA
 51
    DAATPYGEPL PDHFLGAAQM VDELDLLPDA VAATLLADIG RYVPDWNLLV
101 SERCNSTVAE LVKGVDEVQK LTHFARVDSL ATPEERAQQA ETMRKMLLAM
151
201
    VTDIRVVLIK LAMRTRTLQF LSNAPDSPEK RAVAKETLDI FAPLANRLGV
    WOLKWOLEDL GFRHQEPEKY REIALLLDEK RTERLEYIEN FLDILRTELK
    KYNIHFEVAG RPKHIYSIYK KMVKKKLSFD GLFDIRAVRI LVDTVPECYT
301
    TLGIVHSLWQ PIPGEFDDYI ANPKGNGYKS LHTVIVGPEE KGVEVQIRTF
351 DMHQFNEFGV AAHWRYKEGG KGDSAYEQKI AWLRQLLDWR ENMAESGKED
    LAAAFKTELF NDTIYVLTPH GKVLSLPTGA TPIDFAYALH SSIGDRCRGA
401
451
    KVEGQIVPLS TPLENGQRVE IITAKEGHPS VNWLYEGWVK SGKAIGKIRA
    YIRQQNADTV REEGRVQLDK QLAKLTPKPN LQELAENLGY KKPEDLYTAV
551
    GQGEISNRAI QKACGTLNEP PPVPVSATTI VKQSKIKKGG KTGVLIDGED
    GLMTTLAKCC KPAPPDDIAG FVTRERGISV HRKTCPSFRH LAEHAPEKVL
601
    DASWAALQEG QVFAVDIEIR AQDRSGLLRD VSDALARHKL NVTAVQTQSR
651
    DLEASMRFTL EVKQVNDLPR VLAGLGDVKG VLSVTRL*
```

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 419>:

¹ ATGACCGCCA TCAGCCCGAT TCAAGACACG CAAAGCGCGA CTCTGCAAGA

```
51 ATTGCGCGAA TGGTTCGACA GCTACTGCGC CGCTCTGCCG GACAACGATA
      AAAACCTCAT CGGTACCGCA TGGTTGCTGG CGCAGGAACA TTACCCCGCC
 151 GATGCCGCCA CGCCGTATGG CGAGCCGCTG CCCGACCACT TCCTCGGCGC
201 GGCGCAAATG GTTCATGAAC TCGACCTGCT CCCCGATGCC GTCGCCGCCA
      CCCTGCTTGC CGACATCGGA CGCTACGTCC CCGACTGGAA CCTATTGGTT
      TCCGAACGCT GCAACAGTAC CGTCGCCGAG CTGGTCAAAG GTGTGGACGA
      AGTGCAGAAA CTCACCCACT TCGCCCGGGT GGACAGCCTC GCCACGCCGG
 351
      AAGAACGCGC CCAGCAGGCA GAAACTATGC GGAAAATGCT GCTGGCGATG
 401
 451
      GTTACCGACA TCCGCGTCGT GTTAATCAAA CTGGCGATGC GTACGCGCAC
      CCTGCAATTT TTAAGCAACG CCCCGACAG CCCCGAAAAA CGCGCCGTCG
      CCAAAGAAAC CCTCGACATC TTCGCCCCGC TCGCCAACCG TTTGGGCGTG
 551
 601
      TGGCAGCTCA AATGGCAGCT CGAAGATTTG GGCTTCCGCC ATCAAAAGCC
      CGAAAAATAC CGCGAAATCG CGCTGCTTTT GGACGAAAAA CGCACCGAAC
      GCCTCGAATA CATCGAAAAC TTCCTCAACA TCCTGCGCGG TGAACTCAAG
 701
      AAATACAATG TCCATTTCGA AGTCGCCGGC CGCCCGAAAC ACATCTACTC
 801
      CATTTACAAA AAAATGGTGA AGAAAAAACT CAGCTTCGAC GGCCTCTTTG
      ACATCCGCGC CGTGCGAATT CTGGTTGATA CCGTCCCCGA GTGTTACACC
      ACGCTGGGTA TCGTCCACAG CCTCTGGCAG CCCATTCCCG GCGAGTTCGA
 901
      CGACTACATC GCCAATCCCA AAGGCAACGG CTATAAAAGT TTGCACACCG
 951
      TCATCGTCGG CCCGGAAGAC AAAGGCGTGG AAGTACAAAT CCGCACCTTC
1001
      GATATGCACC AATTCAACGA ATTCGGTGTC GCCGCCCACT GGCGTTACAA
1051
      AGAGGGCGC AAGGGCGATT CCGCCTACGA ACAGAAAATC GCCTGGTTGC
1101
      GCCAACTCTT GGACTGGCGC GAAAACATGG CGGAAAGCGG CAAGGAAGAC
1151
1201 CTCGCCGCCG CCTTCAAAAC CGAGCTTTTC AACGACACGA TTTATGTTTT
1251
      GACCCCGCAC GGCAAAGTCC TCTCCCTGCC CACGGGCGCG ACCCCCATCG
1301 ACTTCGCCTA CGCCCTGCAC AGCAGCATCG GCGACCGTTG CCGCGGTGCG
1351 AAAGTCGAAG GGCAGATTGT GCCGCTGTCC ACCCCGCTCG AAAACGGACA
1401 GCGCGTCGAA ATCATTACCG CCAAAGAAGG GCATCCTTCC GTCAACTGGC
1451 TTTACGAAGG CTGGGTCAAA TCCAACAAGG CAATCGGCAA AATCCGCGCC
      TACATCCGCC AGCAAAACGC CGACACCGTG CGCGAAGAAG GCCGCGTCCA
1501
1551 ACTCGACAAA CAGCTTGCCA AACTCACGCC CAAACCCAAC CTGCAAGAGC
1601 TTGCCGAAAA TCTCGGCTAC AAAAAGCCAG AAGACCTCTA CACCGCCGTC 1651 GGACAAGGCG AAATTTCCAA CCGCGCCATC CAAAAAGCCT GCGGCACGCT
1701 GAACGAACCG CCGCCGTAC CCGTCAGCGA AACCACCATC GTCAAACAGT
1751 CCAAAATCAA AAAAGGCGGC AAAAACGGCG TGCTCATCGA CGGCGAAGAC
1801 GGTCTGATGA CCACGCTTGC CAAATGCTGC AAACCCGCGC CGCCCGACGA
1851 TATTATCGGC TTCGTTACCC GCGAGCGCGG CATTTCAGTG CACCGCAAAA
1901
      CCTGCCCGTC TTTCCAACAC CTCGCCGAAC ACGCGCCCGA AAAAGTGCTG
1951 GACGCAAGCT GGGCGGCATT GCAGGAAGGA CAAGTATTCG CCGTCGATAT
2001 CGAAATCCGC GCCCAAGACC GCTCCGGGCT TTTGCGCGAC GTATCCGACG
2051 CGCTCGCCCG CCACAAACTC AACGTTACCG CCGTGCAAAC CCAGTCCCGC
      GACTTGGAAG CCAGCATGAG GTTCACGCTC GAAGTCAAAC AAGTCAACGA
      CCTCCCGCGC GTCCTCGCCA GCCTCGGCGA CGTCAAAGGC GTATTGAGCG
2151
2201 TTACCCGGCT TTAA
```

### This corresponds to the amino acid sequence <SEQ ID 420; ORF 117-1>: m117-1.pep

```
1 MTAISPIQDT QSATLQELRE WFDSYCAALP DNDKNLIGTA WLLAQEHYPA
 51 DAATPYGEPL PDHFLGAAQM VHELDLLPDA VAATLLADIG RYVPDWNLLV
    SERCNSTVAE LVKGVDEVQK LTHFARVDSL ATPEERAQQA ETMRKMLLAM
101
    VTDIRVVLIK LAMRTRTLQF LSNAPDSPEK RAVAKETLDI FAPLANRLGV
201 WQLKWQLEDL GFRHQKPEKY REIALLLDEK RTERLEYIEN FLNILRGELK
    KYNVHFEVAG RPKHIYSIYK KMVKKKLSFD GLFDIRAVRI LVDTVPECYT
301 TLGIVHSLWQ PIPGEFDDYI ANPKGNGYKS LHTVIVGPED KGVEVQIRTF
    DMHQFNEFGV AAHWRYKEGG KGDSAYEQKI AWLRQLLDWR ENMAESGKED
351
401 LAAAFKTELF NDTIYVLTPH GKVLSLPTGA TPIDFAYALH SSIGDRCRGA
451 KVEGQIVPLS TPLENGQRVE IITAKEGHPS VNWLYEGWVK SNKAIGKIRA
501
    YIROQNADTV REEGRVQLDK QLAKLTPKPN LQELAENLGY KKPEDLYTAV
    GOGEISNRAI QKACGTLNEP PPVPVSETTI VKQSKIKKGG KNGVLIDGED
551
    GLMTTLAKCC KPAPPDDIIG FVTRERGISV HRKTCPSFQH LAEHAPEKVL
601
    DASWAALQEG QVFAVDIEIR AQDRSGLLRD VSDALARHKL NVTAVQTQSR
701 DLEASMRFTL EVKQVNDLPR VLASLGDVKG VLSVTRL*
```

ml17-1/g117-1 98.2% identity in 737 aa overlap

	10	20	30	40	50	60
m117-1.pep	MTAISPIQDTQSAT					
		11111111	1111111111	11111111111	1111111111	111111
g117-1	MTAISPIQDTQSAT	LQELREWFDS	YCAALPDNDK	NLIGTAWSLA	QEHYPADAAT	PYGEPL
	10	20	30	40	50	60
	70	80	90	100	110	120

BNSDOCID: <WO___9957280A2_I_>

m117-1.pep	p PDHFLGAAQMVHELDLLPDAVAATLLADIGRYVPDWNLLVSERCNSTVAEI	. W.Cuppuo
g117-1		
,	70 80 90 100 110	LVKGVDEVQK 120
m117-1.pep	130 140 150 160 170	180
		11111111
g117-1	LTHFARVDSLATPEERAQQAETMRKMLLAMVTDIRVVLIKLAMRTRTLQFL 130 140 150 160 170	SNAPDSPEK 180
	190 200 210 220 230	
m117-1.pep	RAVAKETLDIFAPLANRLGVWQLKWQLEDLGFRHQKPEKYREIALLLDEKR	240 TERLEYIEN
g117-1	RAVAKETEDIFAPLANRLGVWQLKWQLEDLGFRHQEPEKYREIALLLDEKR	TERLEYIEN
	250 230	240
m117-1.pep		300 VDTVPECYT
g117-1	:         :	
	250 260 270 280 290	300
m117-1.pep	310 320 330 340 350 TLGIVHSLWQPIPGEFDDYIANPKGNGYKSLHTVIVGPEDKGVEVQIRTFDM	360
q117-1		
9-1. 1	TLGIVHSLWQPIPGEFDDYIANPKGNGYKSLHTVIVGPEEKGVEVQIRTFDN 310 320 330 340 350	MHQFNEFGV 360
110.4	370 380 390 400 410	420
m117-1.pep	AAHWRYKEGGKGDSAYEQKIAWLRQLLDWRENMAESGKEDLAAAFKTELFNI	1111111
g117-1	AAHWRYKEGGKGDSAYEQKIAWLRQLLDWRENMAESGKEDLAAAFKTELFNE 370 380 390 400 410	TIYVLTPH
	420 440 455	420
m117-1.pep	GKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGOTVPLSTPLFMCORVETT	480 TAKEGHPS
g117-1	GKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIVPLSTPLENGQRVEII	 TAKEGHPS
	470	480
m117-1.pep	490 500 510 520 530 VNWLYEGWVKSNKAIGKIRAYIRQQNADTVREEGRVQLDKQLAKLTPKPNLQ	540 ELAENLGY
g117-1		
	490 500 510 520 530	540
m117-1.pep	550 560 570 580 590 KKPEDLYTAVGQGEISNRAIQKACGTLNEPPPVPVSETTIVKQSKIKKGGKNO	600
g117-1		
-	550 560 570 580 590	600
m117-1.pep	610 620 630 640 650	660
g117-1	GLMTTLAKCCKPAPPDDIIGFVTRERGISVHRKTCPSFQHLAEHAPEKVLDAS	
g11 <i>1</i> -1	GLMTTLAKCCKPAPPDDIAGFVTRERGISVHRKTCPSFRHLAEHAPEKVLDAS 610 620 630 640 650	WAALQEG 660
	670 680 690 700 710	720
m117-1.pep	QVFAVDIEIRAQDRSGLLRDVSDALARHKLNVTAVQTQSRDLEASMRFTLEVK	QVNDLPR
g117-1	670 690 COO COO	QVNDLPR
	730	720
m117-1.pep	VLASLGDVKGVLSVTRLX	
g117-1	:	
m117_1/p-:•	730	
m117-1/RelA		

```
sp|P55133|RELA_VIBSS GTP PYROPHOSPHOKINASE (ATP:GTP 3'-PYROPHOSPHOTRANSFERASE) (PPGPP
SYNTHETASE I) >gi|537617 (U13769) ppGpp synthetase I [Vibrio sp.] Length = 744
 Score = 536 bits (1366), Expect = e-151
 Identities = 288/685 (42%), Positives = 432/685 (63%), Gaps = 31/685 (4%)
Query: 74 LDLLPDAVAATLLADI---GRYVPDWNLLVSERCNSTVAELVKGVDEVQKLTHFARVDSL 130
L + D + A LL + G Y D + E + T+ LV+GV+++ ++ S
Sbjct: 68 LSMDADTLIAALLYPLVEGGCYSTD---ALKEEYSGTILHLVQGVEQMCAIS---QLKST 121
Query: 131 ATPEERAQQAETMRKMLLAMVTDIRVVLIKLAMRTRTLQFLSNAPDSPEKRAVAKETLDI 190
              +A Q + +R+MLL+MV D R V+IKLA R L+ + + PD +RA A+E +I
Sbjct: 122 AEETAQAAQVDNIRRMLLSMVDDFRCVVIKLAERICNLREVKDQPDEV-RRAAAQECANI 180
Query: 191 FAPLANRLGVWQLKWQLEDLGFRHQKPEKYREIALLLDEKRTERLEYIENFLNILRGELK 250
           +APLANRLG+ QLKW++ED FR+Q P+ Y++IA L E+R +R +YI +F++ L
Sbjct: 181 YAPLANRLGIGQLKWEIEDYAFRYQHPDTYKQIAKQLSERRIDREDYITHFVDDLSDAMK 240
Query: 251 KYNVHFEVAGRPKHIYSIYKKMVKKKLSFDGLFDIRAVRILVDTVPECYTTLGIVHSLWQ 310
            N+ EV GRPKHIYSI++KM KK L FD LFD+RAVRI+ + + +CY LG+VH+ ++
Sbjct: 241 ASNIRAEVQGRPKHIYSIWRKMQKKSLEFDELFDVRAVRIVAEELQDCYAALGVVHTKYR 300
Query: 311 PIPGEFDDYIANPKGNGYKSLHTVIVGPEDKGVEVQIRTFDMHQFNEFGVAAHWRYKEG- 369
            +P EFDDY+ANPK NGY+S+HTV++GPE K +E+QIRT MH+ +E GVAAHW+YKEG
Sbjct: 301 HLPKEFDDYVANPKPNGYQSIHTVVLGPEGKTIEIQIRTKQMHEESELGVAAHWKYKEGT 360
Query: 370 --GKGDSAYEQKIAWLRQLLDWRENMAESGKEDLAAAFKTELFNDTIYVLTPHGKVLSLP 427
            G SAY++KI WLR+LL W+E M++SG ++ ++++F+D +Y TP G V+ LP
Sbjct: 361 ASGGAQSAYDEKINWLRKLLAWQEEMSDSG--EMLDELRSQVFDDRVYAFTPKGDVVDLP 418
Query: 428 TGATPIDFAYALHSSIGDRCRGAKVEGQIVPLSTPLENGQRVEIITAKEGHPSVNWLYE- 486
           + ATP+DFAY +HS +G RC GAKVEG+IVP + L+ G +VEIIT KE +PS +WL
Sbjct: 419 SNATPLDFAYHIHSEVGHRCIGAKVEGRIVPFTYHLQMGDQVEIITQKEPNPSRDWLNPN 478
Query: 487 -GWVKSNKAIGKIRAYIRQQNADTVREEGRVQLDKQLAKL--TPKPNLQELAENLGYKKP 543
           G+V S++A K+ A+ R+Q+ D G+ L+ +L K+ T K
Sbjct: 479 LGFVTSSRARAKVHAWFRKQDRDKNIIAGKEILEAELVKIHATLKDAQYYAAKRFNVKSP 538
Query: 544 EDLYTAVGQGEIS-NRAIQKACGTLNEPPPVPVSETTIVKQSKI-----KKGGKNGV 594
           E+LY +G G++ N+ I
                              +N+P + + K S+
                                                              KK ++ V
Sbjct: 539 EELYAGIGSGDLRINQVINHINALVNKPTAEEEDQQLLEKLSEASNKQATSHKKPQRDAV 598
Query: 595 LIDGEDGLMTTLAKCCKPAPPDDIIGFVTRERGISVHRKTCPSFQHLAEHAPEKVLDASW 654
           +++G D LMT LA+CC+P P DDI GFVT+ RGISVHR C + L HAPE+++D W
                                                                                      - 30 G
Sbjct: 599 VVEGVDNLMTHLARCCOPIPGDDIQGFVTQGRGISVHRMDCEQLEELRHHAPERIIDTVW 658
Query: 655 AALQEGQVFAVDIEIRAQDRSGLLRDVSDALARHKLNVTAVQTQ--SRDLEASMRFTLEV 712
               G + + + + A +R+GLL+++++ L K+ V ++++ +
                                                             + M F LE+
Sbjct: 659 GGGFVGN-YTITVRVTASERNGLLKELTNTLMNEKVKVAGMKSRVDYKKQMSIMDFELEL 717
Query: 713 KQVNDLPRVLASLGDVKGVLSVTRL 737
             + L RVL + VK V
Sbjct: 718 TDLEVLGRVLKRIEQVKDVAEAKRL 742
```

### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 421>: a117-1.seq

-1.sec	4				
1	ATGACCGCCA	TCAGCCCGAT	TCAAGACACG	CAAAGCGCGA	CTCTGCAAGA
51	ATTGCGCGAA	TGGTTCGACA	GCTACTGCAC	CGCGCTGCCG	AACAACGATA
101	AAAAACTTGT	CTTAGCCGCC	CGTTCGCTGG	CGGAAGCACA	TTACCCCGCC
151	GATGCCGCCA	CGCCGTATGG	CGAACCGCTG	CCCGACCACT	TCCTCGGCGC
201	GGCGCAAATG	GTTCATGAAC	TCGACCTGCT	CCCCGATGCC	GTCGCCGCCA
251	CCCTGCTTGC	CGACATCGGA	CGCTACGTCC	CCGACTGGAA	CCTATTGGTT
301	TCCGAACGCT	GCAACAGTAC	CGTCGCCGAG	CTGGTCAAAG	GTGTGGACGA
351	AGTGCAGAAA	CTCACCCACT	TCGCCCGGGT	GGACAGCCTC	GCCACGCCGG
401	AAGAACGCGC	CCAGCAGGCA	GAAACTATGC	GGAAAATGCT	GCTGGCGATG
451	GTTACCGACA	TCCGCGTCGT	GTTAATCAAA	CTGGCGATGC	GTACGCGCAC
501	CCTGCAATTT	TTAAGCAACG	CCCCGACAG	CCCCGAAAAA	CGCGCCGTCG
551	CCAAAGAAAC	CCTCGACATC	TTCGCCCCGC	TCGCCAACCG	TTTGGGCGTG
601	TGGCAGCTCA	AATGGCAGCT	CGAAGATTTG	GGCTTCCGCC	ATCAAGAACC
651	CGAAAAATAC	CGCGAAATCG	CCCTGCTTTT	GGACGAAAAA	CGCACCGAAC
701	GCCTCGAATA	CATCGAAAAC	TTCCTTAATA	TCCTGCGTAC	GGAACTCAAA
751	AAATACAATA	TCCACTTTGA	AGTCGCCGGC	CGTCCGAAAC	ACATCTACTC
801	CATTTACAAA	AAAATGGTGA	AGAAAAAACT	CAGCTTCGAC	GGGTTGTTCG

```
851 ACATCCGCGC CGTGCGGATT CTGGTTGATA CCGTCCCCGA GTGTTACACC
      ACACTGGGCA TTGTCCACAG CCTCTGGCAG CCCATTCCCG GCGAGTTCGA
  901
      CGACTACATC GCCAACCCGA AAGGCAACGG CTATAAAAGT TTGCACACCG
1001
      TCATCGTCGG CCCGGAAGAC AAAGGCGTGG AAGTGCAAAT CCGCACCTTC
      GATATGCACC AATTCAACGA ATTCGGTGTC GCCGCGCACT GGCGTTACAA
1051
1101 AGAGGGCGGC AAAGGCGATT CCGCCTACGA ACAAAAAATC GCCTGGTTAC
      GCCAACTTTT GGACTGGCGC GAAAACATGG CGGAAAGCGG CAAGGAAGAC
1151
      CTCGCCGCCG CCTTCAAAAC CGAGCTTTTC AACGACACGA TTTATGTTTT
1201
1251
      GACCCCGCAC GGCAAAGTCC TCTCCCTGCC CACAGGCGCG ACCCCCATCG
1301 ACTTCGCCTA CGCCCTGCAC AGCAGCATCG GCGACCGTTG CCGCGGTGCG
1351 AAAGTCGAAG GGCAGATTGT GCCGCTGTCC ACCCCGCTCG AAAACGGACA
      GCGTGTCGAA ATCATTACCG CCAAAGAAGG GCATCCTTCC GTCAACTGGC
1401
1451
      TTTACGAAGG CTGGGTCAAA TCCAACAAGG CAATCGGCAA AATCCGCGCC
1501
      TACATCCGCC AGCAAAACGC CGACACCGTG CGCGAAGAAG GCCGCGTCCA
1551 ACTCGACAAA CAGCTTGCCA AACTCACGCC CAAACCCAAC CTGCAAGAGC
      TTGCCGAAAA TCTCGGCTAC AAAAAGCCAG AAGACCTCTA CACCGCCGTC
1601
1651 GGACAAGGCG AAATTTCCAA CCGCGCCATC CAAAAAGCCT GCGGCACGCT
1701 GAACGAACCG CCGCCCGTAC CCGTCAGCGA AACCACCATC GTCAAACAGT
1751 CCAAAATCAA AAAAGGCGGC AAAAACGGCG TGCTCATCGA CGGCGAAGAC
1801 GGTCTGATGA CCACGCTTGC CAAATGCTGC AAACCCGCGC CGCCCGACGA
     CATTGTCGGC TTCGTTACCC GCGATCGCGG CATTTCGGTA CACCGCAAAA
1851
1901
     CCTGCCCCTC TTTCCGACAC CTCGCCGAAC ACGCGCCCGA AAAAGTACTG
1951
     GACGCAAGTT GGGCGGCGTT GCAGGAAGGA CAAGTGTTCG CCGTCGATAT
     CGAAATCCGC GCCCAAGACC GCTCCGGGCT TTTGCGCGAC GTATCCGACG
2001
     CGCTCGCCCG CCACAAACTC AACGTTACCG CCGTGCAAAC CCAGTCCCGC
2051
     GACTTGGAAG CCAGCATGAG GTTCACGCTC GAAGTCAAAC AAGTTACCGA
2101
     CCTCCCACGC GTCCTCGCCA GCCTCGGCGA CGTCAAAGGC GTATTGAGCG
2151
2201 TTACCCGGCT TTAA
```

# This corresponds to the amino acid sequence <SEQ ID 422; ORF 117-1.a>:

```
1 MTAISPIQDT QSATLQELRE WFDSYCTALP NNDKKLVLAA RSLAEAHYPA
 51 DAATPYGEPL PDHFLGAAQM VHELDLLPDA VAATLLADIG RYVPDWNLLV
101
     SERCNSTVAE LVKGVDEVQK LTHFARVDSL ATPEERAQQA ETMRKMLLAM
151 VTDIRVVLIK LAMRTRTLQF LSNAPDSPEK RAVAKETLDI FAPLANRLGV
201 WQLKWQLEDL GFRHQEPEKY REIALLLDEK RTERLEYIEN FLNILRTELK
251 KYNIHFEVAG RPKHIYSIYK KMVKKKLSFD GLFDIRAVRI LVDTVPECYT
301 TLGIVHSLWQ PIPGEFDDYI ANPKGNGYKS LHTVIVGPED KGVEVQIRTF
    DMHQFNEFGV AAHWRYKEGG KGDSAYEQKI AWLRQLLDWR ENMAESGKED
351
401
    LAAAFKTELF NDTIYVLTPH GKVLSLPTGA TPIDFAYALH SSIGDRCRGA
    KVEGQIVPLS TPLENGQRVE IITAKEGHPS VNWLYEGWVK SNKAIGKIRA
451
    YIRQQNADTV REEGRVQLDK QLAKLTPKPN LQELAENLGY KKPEDLYTAV
501
    GQGEISNRAI QKACGTLNEP PPVPVSETTI VKQSKIKKGG KNGVLIDGED
551
    GLMTTLAKCC KPAPPDDIVG FVTRDRGISV HRKTCPSFRH LAEHAPEKVL
    DASWAALQEG QVFAVDIEIR AQDRSGLLRD VSDALARHKL NVTAVQTQSR
651
    DLEASMRFTL EVKQVTDLPR VLASLGDVKG VLSVTRL*
```

### a117-1/m117-1 97.7% identity in 737 aa overlap

	10	20	30	40	50	60
m117-1.pep	MTAISPIQDTQSAT	LQELREWFDS	SYCAALPONDE	KNLIGTAWLL	AOEHYPADAA	*******
			1 [ :   ]   :   ]	1.1	1. 1111111	
a117-1	MTAISPIQDTQSAT	LQELREWFDS	YCTALPNND	KLVLAARSI	1 - 1	POVCEDI
	10	20	30	40	50	
				40	30	60
	70	80	90	100	110	120
m117-1.pep	PDHFLGAAQMVHEI	DLLPDAVAAT	LLADIGRYVE	DWNLLVSER	אוכידעו די נוער	TIDEUON
		111111111	111111111			
a117-1	PDHFLGAAQMVHEL	DLLPDAVAAT	LLADICRYVE	ODWNITUREDO	.	11111
	70	80	90	100		
		••	50	100	110	120
	130	140	150	•		
m117-1.pep			150	160	170	180
	LTHFARVDSLATPE	ERAQQALIMR	KMLLAMVTDI	RVVLIKLAMF	(TRTLQFLSNA	PDSPEK
a117-1			11111111	111111111	1111111111	111111
a11/-1	DIREARVOSLATPE	EKAQQAETMR	KMLLAMVTDI	RVVLIKLAME	TRTLOFLSNA	PDSPEK
	130	140	150	160	170	180
	190	200	210	220	230	240
m117-1.pep	RAVAKETLDIFAPL	ANRLGVWQLK	WQLEDLGFRH	OKPEKYRETA	TIIDDUDANN	
	- 1 1 1 2 1 1 2 1 4 1 4 1 1 1 1 1 1 1	1 1 1 1 1 1 1 1 1 1	111111111	F • 1 1 1 1 1 1 1 1 1		
al17-1	RAVAKETLDIFAPL	ANRLGVWOLK	WOLEDI GERH	OEDEKVDETA	III DEVENDED	11111
	190	200	210	220		
			~ ~ 0	220	230	240

m117-1.pep	250 260 270 280 290 300  FLNILRGELKKYNVHFEVAGRPKHIYSIYKKMVKKKLSFDGLFDIRAVRILVDTVPECYT
m117-1.pep	310 320 330 340 350 360 TLGIVHSLWQPIPGEFDDYIANPKGNGYKSLHTVIVGPEDKGVEVQIRTFDMHQFNEFGV
m117-1.pep	370 380 390 400 410 420 AAHWRYKEGGKGDSAYEQKIAWLRQLLDWRENMAESGKEDLAAAFKTELFNDTIYVLTPH !!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
m117-1.pep	430 440 450 460 470 480 GKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIVPLSTPLENGQRVEIITAKEGHPS
m117-1.pep	490 500 510 520 530 540 VNWLYEGWVKSNKAIGKIRAYIRQQNADTVREEGRVQLDKQLAKLTPKPNLQELAENLGY
ml17-1.pep	550 560 570 580 590 600  KKPEDLYTAVGQGEISNRAIQKACGTLNEPPPVPVSETTIVKQSKIKKGGKNGVLIDGED
m117-1.pep	610 620 630 640 650 660 GLMTTLAKCCKPAPPDDIIGFVTRERGISVHRKTCPSFQHLAEHAPEKVLDASWAALQEG
m117-1.pep	670 680 690 700 710 720 QVFAVDIEIRAQDRSGLIRDVSDALARHKLNVTAVQTQSRDLEASMRFTLEVKQVNDLPR
ml17-1.pep	730 VLASLGDVKGVLSVTRLX

### The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 423>: g118.seq

ATGTGCGAGT TCAAGGATTT TAGAAGAAAC ATCCCTTGTT TTGAAGAGTA 51 TGACGAAAAT TCATTTATTG GCAAATGGTA TGATGACGGG GTGTGGGATG 101 ATGAAGAATA TTGGAAGCTG GAGAATGATT TAATCGAGGT TAGGAGAAAA 151 TATCCTTATC CGATGGATAT ACCAAGGGAT ATTGTGATTG GAATCGGTAC 201 CATTATTGAT TTTTTAATGG TTCCAAATTG GGAGCTTTTT GAAATTAAAG 251 CTTCCCCTTG GTTGCCTGAT AGCGTGGGAA TTCATGAACG TTATGAAAGA 301 TTCACAACGA TGCTCCGTTA TATTTTTACC GAGAAAGACA TAGTCAACGT 351 GCGATTTGAT TATTACAaCA AAAAATAG

This corresponds to the amino acid sequence <SEQ ID 424; ORF 118.ng>:

- 1 MCEFKDFRRN IPCFEEYDEN SFIGKWYDDG VWDDEEYWKL ENDLIEVRRK
- 51 YPYPMDIPRD <u>IVIGIGTIID FLMVPNW</u>ELF EIKASPWLPD SVGIHERYER 101 FTTMLRYIFT <u>EKDIVNVRFD YYNKK*</u>

BNSDOCID: <WO___9957280A2_I_>

```
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 425>:
      ml18.seq
                ATGTGTGAGT TCAAGGATAT TATAAGAAAC GTTCCTTATT TTGAGGGGTA
                TGACGAAAAT TCATTTATTG GCAAATGGTA TGATGACGGG GTGTGGGATG
            51
           101
                ATGAAGAATA TTGGAAGTTG GAGAATGATT TAATCGAGGT TAGAAAAAAA
                TATCCTTATC CGATGGACAT ACCAAGATAT GTTGTCATTG GAATCGGTAC
           151
               CATTATTGAT TTCTTAATGG TTCCAAATTG GAAACTTTTT GAAATTAAAG
           201
               CTTCCCCTTG GTTGCCTGAT AGTGTGGGAA TTCATGAACG TTATGAAAGA
                TTCACAACGA TGCTCCGTTA TATTTTTACC GAGAAAGACA TAGTCAACGT
           301
           351
                GCGATTTGAT TATTACAACA AAAAATAG
 This corresponds to the amino acid sequence <SEQ ID 426; ORF 118>:
      m118.pep
                MCEFKDIIRN VPYFEGYDEN SFIGKWYDDG VWDDEEYWKL ENDLIEVRKK
                YPYPMDIPRY <u>VVIGIGTIID FLMVPNW</u>KLF EIKASPWLPD SVGIHERYER FTTMLRYIFT EKDIVNVRFD YYNKK*
 Computer analysis of this amino acid sequence gave the following results:
 Homology with a predicted ORF from N. gonorrhoeae
 ORF 118 shows 92.8% identity over a 125 aa overlap with a predicted ORF (ORF 118.ng)
 from N. gonorrhoeae:
      m118/g118
                          10
                                    20
                                             30
                                                       40
                                                                50
                  MCEFKDIIRNVPYFEGYDENSFIGKWYDDGVWDDEEYWKLENDLIEVRKKYPYPMDIPRY
      m118.pep
                  MCEFKDFRRNIPCFEEYDENSFIGKWYDDGVWDDEEYWKLENDLIEVRRKYPYPMDIPRD
      g118
                          10
                                   20
                                             30
                                                       40
                                                                50
                                   80
                                             90
                                                      100
                                                               110
                  VVIGIGTIIDFLMVPNWKLFEIKASPWLPDSVGIHERYERFTTMLRYIFTEKDIVNVRFD
     m118.pep
                  q118
                  IVIGIGTIIDFLMVPNWELFEIKASPWLPDSVGIHERYERFTTMLRYIFTEKDIVNVRFD
                                   80
                                             90
                                                     100
                                                               110
     m118.pep
                  YYNKKX
                  111111
     g118
                  YYNKKX
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 427>:
a118.seq
               ATGTGTGAGT TCAAGGATTT TAGAAGAAAC ATCCCTTGTT TTGAAGAGTA
           51
               TGACGAAAAT TCATTTATTG GCAAATGGTA TGATGACGGG GTGTGGGATG
          101 ATGAAGAATA TTGGAAATTG GAGAATGATT TAATCGAGGT TAGAAAAAAA
               TATCCTTATC CGATGGATAT ACCAAGGGAT ATTGTGATTG GAATCGGTAC
          201
               CATTATTGAT TTTTTAATGG TTCCAAATTG GGAGCTTTTT GAAATTAAAG
          251
               CTTCCCCTTG GTTGCCTGAT AGTGTGGGAA TTCATGAACG TTATGAAAGA
          301
               TTCACAACGA TGCTCCGTTA TATTTTTACC GAGAAAGACA TAGTCAACGT
               GCGATTTGAT TATTACAACA AAAAATAG
This corresponds to the amino acid sequence <SEQ ID 428; ORF 118.a>:
a118.pep
            1 MCEFKDFRRN IPCFEEYDEN SFIGKWYDDG VWDDEEYWKL ENDLIEVRKK
              YPYPMDIPRD IVIGIGTIID FLMVPNWELF EIKASPWLPD SVGIHERYER
          101 FTTMLRYIFT EKDIVNVRFD YYNKK*
m118/a118
              93.6% identity in 125 aa overlap
                                  20
                                            30
                                                      40
     m118.pep
                 MCEFKDIIRNVPYFEGYDENSFIGKWYDDGVWDDEEYWKLENDLIEVRKKYPYPMDIPRY
                 a118
                 MCEFKDFRRNIPCFEEYDENSFIGKWYDDGVWDDEEYWKLENDLIEVRKKYPYPMDIPRD
                                  20
                                            30
                                                     40
                                                               50
```

. :.cr.-g

```
90
                                       100
                                               110
                                                       120
          VVIGIGTIIDFLMVPNWKLFEIKASPWLPDSVGIHERYERFTTMLRYIFTEKDIVNVRFD
m118.pep
          a118
          IVIGIGTIIDFLMVPNWELFEIKASPWLPDSVGIHERYERFTTMLRYIFTEKDIVNVRFD
                70
                        80
                                90
                                       100
m118.pep
          YYNKKX
          11111
a118
          YYNKKX
```

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 429>:

```
g120.seq
         ATGATGAAGA CTTTTAAAAA TATATTTTCC GCCGCCATTT TGTCCGCCGC
      51
         CCTGCCGTGC GCGTATGCGG CAAGGCTACC CCAATCCGCC GTGCTGCACT
          ATTCCGGCAG CTACGGCATT CCCGCCACGA TGACATTTGA ACGCAGCGGC
     101
     151
         AATGCTTACA AAATCGTTTC GACGATTAAA GTGCCGCTAT ACAATATCCG
     201 TTTCGAATCC GGCGGTACGG TTGTCGGCAA TACCCTGCAC CCTGCCTACT
     251 ATAAAGACAT ACGCAGGGGC AAACTGTATG CGGAAGCCAA ATTCGCCGAC
     301 GGCAGCGTAA CCTACGGCAA AGCGGGCGAG AGCAAAACCG AGCAAAGCCC
     351 CAAGGCTATG GATTTGTTCA CGCTTGCCTG GCAGTTGGCG GCAAATGACG
     401 CGAAACTCCC CCCGGGTCTG AAAATCACCA ACGGCAAAAA ACTTTATTCC
     451 GTCGGCGGCC TGAATAAGGC GGGTACGGGA AAATACAGCA Taggcggcqt
     501 GGAAACCGAA GTCGTCAAAT ATCGGGTGCG GCGCGGCGAC GATACGGTAA
     551 CGTATTTCTT CGCACCGTCC CTGAACAATA TTCCGGCACA AATCGGCTAT
     601 ACCGACGACG GCAAAACCTA TACGCTGAAG CTCAAATCGG TGCAGATCAA
         CGGACAGCC GCCAAACCGT AA
```

This corresponds to the amino acid sequence <SEO ID 430; ORF 120.ng>: g120.pep

MMKTFKNIFS AAILSAALPC AYAARLPQSA VLHYSGSYGI PATMTFERSG 51 NAYKIVSTIK VPLYNIRFES GGTVVGNTLH PAYYKDIRRG KLYAEAKFAD 101 GSVTYGKAGE SKTEQSPKAM DLFTLAWQLA ANDAKLPPGL KITNGKKLYS VGGLNKAGTG KYSIGGVETE VVKYRVRRGD DTVTYFFAPS LNNIPAOIGY 201 TDDGKTYTLK LKSVQINGQA AKP*

The following partial DNA sequence was identified in N. meningitidis <SEO ID 431>: m120.seq

ATGATGAAGA CTTTTAAAAA TATATTTTCC GCCGCCATTT TGTCCGCCGC 1 CCTGCCGTGC GCGTATGCGG CAGGGCTGCC CCAATCCGCC GTGCTGmACT 101 ATTCCGGCAG CTACGGCATT CCCGCCACGA TGACATTTGA ACGCAGCGGC 151 AATGCTTACA AAATCGTTTC GACGATTAAA GTGCCGCTAT ACAATATCCG 201 TTTCGAGTCC GGCGGTACGG TTGTCGGCAA TACCCTGCAC CCTACCTACT ATAGAGACAT ACGCAGGGGC AAACTGTATG CGGAAGCCAA ATTCGCCGAC GGCAGCGTAA CTTACGGCAA AGCGGGCGAG AGCAAAACCG AGCAAAGCCC CAAGGCTATG GATTTGTTCA CGCTTGCCTG GCAGTTGGCG GCAAATGACG CGAAACTCCC CCCGGGGCTG AAAATCACCA ACGGCAAAAA ACTTTATTCC 451 GTCGGCGGTT TGAATAAGGC GGGTACAGGA AAATACAGCA TAGGCGGCGT 501 GGAAACCGAA GTCGTCAAAT ATCGGGTGCG GCGCGGCGAC GATGCGGTAA

551 TGTATTTCTT CGCACCGTCC CTGAACAATA TTCCGGCACA AATCGGCTAT 601 ACCGACGACG GCAAAACCTA TACGCTGAAA CTCAAATCGG TGCAGATCAA

CGGCCAGGCA GCCAAACCG

This corresponds to the amino acid sequence <SEQ ID 432; ORF 120>: m120.pep

> MMKTFKNIFS AAILSAALPC AYAAGLPQSA VLXYSGSYGI PATMTFERSG 51 NAYKIVSTIK VPLYNIRFES GGTVVGNTLH PTYYRDIRRG KLYAEAKFAD 101 GSVTYGKAGE SKTEQSPKAM DLFTLAWQLA ANDAKLPPGL KITNGKKLYS 151 VGGLNKAGTG KYSIGGVETE VVKYRVRRGD DAVMYFFAPS LNNIPAOIGY

201 TDDGKTYTLK LKSVQINGQA AKP

Computer analysis of this amino acid sequence gave the following results: Homology with a predicted ORF from N. gonorrhoeae

ORF 120 shows 97.3% identity over a 223 aa overlap with a predicted ORF (ORF 120.ng) from N. gonorrhoeae:

m120/g120

m120.pep	10 MMKTFKNIFSAAII	20 SAALPCAYAA	30 AGLPQSAVLX	40 (SGSYGIPATM	50 ITFERSGNAYI	60 CIVSTIK
g120			 RLPQSAVLHY			 IVSTIK
	10	20	30	40	50	60
m120.pep	70	. 80	90	100	110	120
220.pcp	VPLYNIRFESGGTV	VGNILHPIYY	RDIRRGKLYA	EAKFADGSVT	YGKAGESKTE	QSPKAM
		11111:11	: [ ] ] ] ] ] ] ]		1111111111	111111
g120	VPLYNIRFESGGTV	VGNTLHPAYY	KDIRRGKLYA	EAKFADGSVT	YGKAGESKTE	OSDKVM
	70	80	90	100	110	120
				200	110	120
	130	140	150	7.00		
m120.pep				160	170	180
bcb	DLFTLAWQLAANDA	VIDEGRATIN	GKKTASACGT	NKAGTGKYSI	GGVETEVVKY	RVRRGD
			111111111			
g120	DLFTLAWQLAANDA	KLPPGLKITN	GKKLYSVGGL	NKAGTGKYSI	GGVETEVVKY	RVRRGD
	130	140	150	160	170	180
					170	180
	190	200	210	220	-	
m120.pep	DAVMYFFAPSLNNI:	PAQIGYTDDG	KTYTLKLKSV	OINGOAAKP		
	1:1 111111111	111111111				
g120	DTVTYFFAPSLNNI	PACTGVTDDG	!	OTMOODER		
~	190					
	. 190	200	210	220		

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 433>: a120.seq

```
ATGATGAAGA CTTTTAAAAA TATATTTTCC GCCGCCATTT TGTCCGCCGC
     CCTGCCGTGC GCGTATGCGG CAGGGCTGCC CCAATCCGCC GTGCTGCACT
 51
101 ATTCCGGCAG CTACGGCATT CCCGCCACGA TGACATTTGA ACGCAGCGGC
151 AATGCTTACA AAATCGTTTC GACGATTAAA GTGCCGCTAT ACAATATCCG
     TTTCGAGTCC GGCGGTACGG TTGTCGGCAA TACCCTGCAC CCTACCTACT
201
251
    ATAGAGACAT ACGCAGGGGC AAACTGTATG CGGAAGCCAA ATTCGCCGAC
     GGCAGCGTAA CCTACGGCAA AGCGGGCGAG AGCAAAACCG AGCAAAGCCC
301
351 CAAGGCTATG GATTTGTTCA CGCTTGCCTG GCAGTTGGCG GCAAATGACG
401 CGAAACTCCC CCCGGGGCTG AAAATCACCA ACGGCAAAAA ACTTTATTCC
451 GTCGGCGGTT TGAATAAGGC GGGTACAGGA AAATACAGCA TAGGCGGCGT
    GGAAACCGAA GTCGTCAAAT ATCGGGTGCG GCGCGGCGAC GATGCGGTAA
    TGTATTTCTT CGCACCGTCC CTGAACAATA TTCCGGCACA AATCGGCTAT
551
    ACCGACGACG GCAAAACCTA TACGCTGAAA CTCAAATCGG TGCAGATCAA
601
651
    CGGCCAGGCA GCCAAACCGT AA
```

This corresponds to the amino acid sequence <SEQ ID 434; ORF 120.a>: a120.pep

```
1 MMKTFKNIFS AAILSAALPC AYAAGLPQSA VLHYSGSYGI PATMTFERSG
51 NAYKIVSTIK VPLYNIRFES GGTVVGNTLH PTYYRDIRRG KLYAEAKFAD
101 GSVTYGKAGE SKTEQSPKAM DLFTLAWQLA ANDAKLPPGL KITNGKKLYS
151 VGGLNKAGTG KYSIGGVETE VVKYRVRRGD DAVMYFFAPS LNNIPAQIGY
201 TDDGKTYTLK LKSVQINGQA AKP*
```

21.0 (21.001.11.11

m120/a120 99.6% identity in 223 aa overlap

m120.pep	10	20	30	40	50	60
mizo.beb	MMKTFKNIFSAAIL.	SAALPCAYAA	GLPQSAVLXYS	GSYGIPATM	TFERSGNAYK	IVSTIK
a120		111111111	11111111		1111111111	Litia

	10	20	30	40	50	60
	70	80	90	100	110	120
m120.pep	VPLYNIRFESGGT	VVGNTLHPTYY	RDIRRGKLYA	EAKFADGSVI	TYGKAGESKT	EQSPKAM
		11111111111		1111111111		HILLE
a120	VPLYNIRFESGGT	'VVGNTLHPTYY	RDIRRGKLYA	EAKFADGSVI	YGKAGESKT	EQSPKAM
	70	80	90	100	110	120
	130	140	150	160	170	180
m120.pep	DLFTLAWQLAAND	AKLPPGLKITN	IGKKLYSVGGL	NKAGTGKYSI	GGVETEVVKY	RVRRGD
		1111111111	1111111111		111111111	
a120	DLFTLAWQLAAND		GKKLYSVGGL	nkagtgkysi	GGVETEVVKY	RVRRGD
	130	140	150	160	170	180
	190	200	210	220		
m120.pep	DAVMYFFAPSLNN	IPAQIGYTDDG	KTYTLKLKSV	QINGQAAKPX		
• •		1111111111	1111111111	ÎHILIHH		
a120	DAVMYFFAPSLNN	IPAQIGYTDDG	KTYTLKLKSV	QINGQAAKPX		
	190	200	210	220		

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 435>: g121.seq

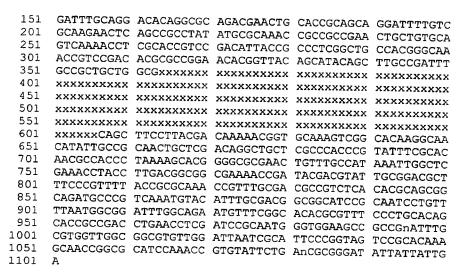
```
ATGGAAACAC AGCTTTACAT CGGCATTATG TCGGGAACCA GTATGGACGG
      GGCGGATGCC GTGCTGGTAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
  51
     AAGGGCACGC CTTTACCCCC TACCCTGACC GGTTGCGCCG CAAATTGCTG
 101
 151 GATTTGCAGG ACACAGGCAC AGACGAACTG CACCGCAGCA GGATGTTGTC
 201 GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
 251
     GTCAAAACCT CGCTCCGTGC GACATTACCG CCCTCGGCTG CCACGGGCAA
     ACCGTCCGAC ACGCGCCGGA ACACGGTtac AGCATACAGC TTGCCGATTT
 301
 351
      GCCGCTGCTG GCGGAACTGa cgcggatttT TACCGTCggc gacttcCGCA
 401 GCCGCGACCT TGCTGCCGGC GGacaAGGTG CGCCGCTCGT CCCCGCCTTT
     CACGAAGCCC TGTTCCGCGA TGACAGGGAA ACACGCGTGG TACTGAACAT
 451
 501 CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCGGCGCA CCCGCCTTCG
 551 GCTTCGACAC AGGGCCGGGC AATATGCTGA TGGAcgcgtg gacgcaggca
     cacTGGcagc TGCCTTACGA CAAAAacggt gcAAAGgcgg cacAAGGCAA catatTGCcg cAACTGCTCG gcaggctGCT CGCCcaccCG TATTTCTCAC
 601
 651
 701 AACCCcaccc aaAAAGCACG GGgcGCGaac TgtttgcccT AAattggctc
 751 gaaacctAcc ttgacggcgg cgaaaaccga tacgacgtat tgcggacgct
801 ttcccgattc accgcgcaaA ccgTttggga cgccgtctca CACGCAGCGG
 851
     CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901
      TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAA CTGAACCTCG ATCCTCAATG GGTGGAGGCG gccgCATTtg
1001
     cgtggttggC GGCGTGTTGG ATTAACCGCA TTCCCGGTAG TCCGCACAAA
1051
     GCGACCGGCG CATCCAAACC GTGTATTCTG GGCGCGGGAT ATTATTATTG
1101 A
```

This corresponds to the amino acid sequence <SEQ ID 436; ORF 121.ng>: g121.pep

1	METQLYIGIM	SGTSMDGADA	VLVRMDGGKW	LGAEGHAFTP	YPDRLRRKLL
51	DLQDTGTDEL	HRSRMLSQEL	SRLYAQTAAE	LLCSQNLAPC	DITALGCHGQ
101	TVRHAPEHGY	SIQLADLPLL	AELTRIFTVG	DFRSRDLAAG	GQGAPLVPAF
151	HEALFRDDRE	TRVVLNIGGI	ANISVLPPGA	PAFGFDTGPG	NMLMDAWTQA
201	HWQLPYDKNG	AKAAQGNILP	QLLGRLLAHP	YFSQPHPKST	GRELFALNWL
251	ETYLDGGENR	YDVLRTLSRF	TAQTVWDAVS	HAAADARQMY	ICGGGIRNPV
301	LMADLAECFG	TRVSLHSTAE	LNLDPQWVEA	AAFAWLAACW	INRIPGSPHK
351	ATGASKPCIL	GAGYYY*			

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 437>: m121.seq

1	ATGGAAACAC	AGCTTTACAT	CGGCATCATG	TCGGGAACCA	GCATGGACGG
51	GGCGGATGCC	GTACTGATAC	GGATGGACGG	CGGCAAATGG	CTGGGCGCGG
101	AAGGGCACGC	CTTTACCCCC	TACCCCGGCA	GGTTACGCCG	CCAATTGCTG



This corresponds to the amino acid sequence <SEQ ID 438; ORF 121>: m121.pep

1	METQLYIGIM	SGTSMDGADA	VLIRMDGGKW	LGAEGHAFTP	YPGRT.PROTT
51	DLQDTGADEL	HRSRILSQEL	SRLYAQTAAE	LLCSONLAPS	DITALCCHOO
101	TVRHAPEHGY	SIQLADLPLL	Axxxxxxxx	XXXXXXXXXX	STINDGCUGO
151	XXXXXXXXX	xxxxxxxxx	xxxxxxxxx	XXXXXXXXX	XXXXXXXXXX
201	XXQLPYDKNG	AKSAQGNILP	QLLDRLLAHP	YFAORHPKST	GRELEATNWI
251	ETYLDGGENR	YDVLRTLSRF	TAQTVCDAVS	HAAADAROMY	TCDGGTRNPV
301	LMADLAECFG	TRVSLHSTAD	LNLDPQWVEA	AXFAWLAACW	INRIPGSPHK
351	ATGASKPCIL	XAGY <u>YY</u> *			

Computer analysis of the amino acid sequences gave the following results: Homology with a predicted ORF from N. meningitidis menA with menB ORF 121 shows 73.5% identity over a 366 aa overlap with a predicted ORF (ORF121.ng) from N. gonorrhoeae:

m121/g121

	10	20	30	40	50	60
m121.pep	METQLYIGIMSGTS	MDGADAVLI:	RMDGGKWLGAR	GHAFTPYPO	ERT.RROLL DIO	סט פט
		11111111:	1111111111			
g121	METQLYIGIMSGTS	MDGADAVIV	RMDGGKWI.GAF	CHAPTOVDE		111:111
	10	20	30	40		
	70	80	90		50	60
m121.pep			90 20NT 7 DOD TM7	100	110	120
PP	HRSRILSQELSRLY	VOINATITO	PONTAPSOTIA	LECHGOTVE	RHAPEHGYSIQ:	LADLPLL
g121				11111111	111111111	111111
9121	HRSRMLSQELSRLY	AQTAAELLC	SQNLAPCDITA	LGCHGQTVF	HAPEHGYSIQ:	LADLPLL
	70	80	90	100	110	120
101	130	140	150	160	170	180
m121.pep	AXXXXXXXXXXXXX	XXXXXXXXX	XXXXXXXXXXX	XXXXXXXX	XXXXXXXXXX	XXXXXX
	1 : :			•		
g121	AELTRIFTVGDFRS	RDLAAGGQGA	APLVPAFHEAL	FRDDRETRV	VINTEGTANT	משמת זונים
	130	140	150	160	170	
	190	200	210	220	230	180
m121.pep	XXXXXXXXXXXXXX	XXXXXXX	PADRNCVRCV	OCNIT DOLL	230	240
	•		ILLIIIIII	ZGMITE ČET		
g121	PAFGEDTGPGNMIM		DYDINGRUE	111111111		11111
3	PAFGFDTGPGNMLM 190	200	PIDKNGAKAA	QGNILPQLL		PHPKST
	250		210	220	230	240
m121.pep		260	270	280	290	300
mrzi.pep	GRELFAINWLETYL	DGGENRYDVI	RTLSRFTAQT	VCDAVSHAA	ADARQMYICDG	GIRNPV
-101		;		1 111111	I FERRITE F	factor of
g121	GRELFALNWLETYL	DGGENRYDVL	RTLSRFTAQT	VWDAVSHAA	ADAROMYICGG	GTRNIPU
	250	260	270	280	290	300

	310	320	330	340	350	360
m121.pep	LMADLAECFGTRV	SLHSTADLNL	DPQWVEAAXFA	WLAACWINR	PGSPHKATG	ASKPCIL
		1111111:111		111111111		
g121	LMADLAECFGTRV	SLHSTAELNL	DPQWVEAAAFA	WLAACWINR	PGSPHKATG	ASKPCIL
•	310	320	330	340	350	360
m121.pep	XAGYYYX					
g121	GAGYYYX					

#### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 439>:

```
al21.seq
          ATGGAAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
          GGCGGATGCC GTACTGATAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
      51
         AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTTACGCCG CAAATTGCTG
     101
          GATTTGCAGG ACACAGGCGC GGACGAACTG CACCGCAGCA GGATGTTGTC
     151
          GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
     201
     251
          GTCAAAACCT CGCGCCGTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
         ACCGTCAGAC ACGCGCCGGA ACACAGTTAC AGCGTACAGC TTGCCGATTT
     301
          GCCGCTGCTG GCGGAACGGA CTCAGATTTT TACCGTCGGC GACTTCCGCA
     351
          GCCGCGACCT TGCGGCCGGC GGACAAGGCG CGCCGCTCGT CCCCGCCTTT
         CACGAAGCCC TGTTCCGCGA CGACAGGGAA ACACGCGCGG TACTGAACAT
     451
          CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCGACGCA CCCGCCTTCG
     501
     551
          GCTTCGACAC AGGACCGGGC AATATGCTGA TGGACGCGTG GATGCAGGCA
          CACTGGCAGC TTCCTTACGA CAAAAACGGT GCAAAGGCGG CACAAGGCAA
     601
          CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCCACCCG TATTTCGCAC
     651
          AACCCCACCC TAAAAGCACG GGGCGCGAAC TGTTTGCCCT AAATTGGCTC
          GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TGCGGACGCT
     751
     801
          TTCCCGATTC ACCGCGCAAA CCGTTTTCGA CGCCGTCTCA CACGCAGCGG
     851
          CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
         TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
     901
          CACCGCCGAA CTGAACCTCG ATCCGCAATG GGTAGAAGCC GCCGCGTTCG
     951
    1001
         CATGGATGGC GGCGTGTTGG GTCAACCGCA TTCCCGGTAG TCCGCACAAA
    1051
         GCAACCGGCG CATCCAAACC GTGTATTCTG GGCGCGGGAT ATTATTATTG
    1101
```

#### This corresponds to the amino acid sequence <SEQ ID 440; ORF 121.a>:

```
a121.pep

1 METQLYIGIM SGTSMDGADA VLIRMDGGKW LGAEGHAFTP YPGRLRRKLL
51 DLQDTGADEL HRSRMLSQEL SRLYAQTAAE LLCSQNLAPS DITALGCHGQ
101 TVRHAPEHSY SVQLADLPLL AERTQIFTVG DFRSRDLAAG GQGAPLVPAF
151 HEALFRDDRE TRAVLNIGGI ANISVLPPDA PAFGFDTGPG NMLMDAWMQA
201 HWQLPYDKNG AKAAQGNILP QLLDRLLAHP YFAQPHPKST GRELFALNWL
251 ETYLDGGENR YDVLRTLSRF TAQTVFDAVS HAAADARQMY ICGGGIRNPV
301 LMADLAECFG TRVSLHSTAE LNLDPQWVEA AAFAWMAACW VNRIPGSPHK
351 ATGASKPCIL GAGYYY*
```

#### m121/a121 74.0% identity in 366 aa overlap

```
10
                       20
                               30
                                       40
                                              50
          METQLYIGIMSGTSMDGADAVLIRMDGGKWLGAEGHAFTPYPGRLRRQLLDLQDTGADEL
m121.pep
          METQLYIGIMSGTSMDGADAVLIRMDGGKWLGAEGHAFTPYPGRLRRKLLDLQDTGADEL
a121
                10
                        20
                               30
                                       40
                                              50
                                                      60
                        80
                               90
                70
                                      100
                                             110
          HRSRILSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPEHGYSIQLADLPLL
m121.pep
          a121
          HRSRMLSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPEHSYSVQLADLPLL
                70
                       80
                               90
                                      100
```

BNSDOCID: <WO__9957280A2_l_>

m121.pep	130 AXXXXXXXXXXXXXXXX	140 XXXXXXXXX	150 XXXXXXXXX	160 XXXXXXXXXX	170 XXXXXXXXX	180 XXXXXX
a121	AERTQIFTVGDFR:	SRDLAAGGQG	APLVPAFHEA	: LFRDDRETRAN	T.NTGGTANT	מתפם זעס
	130	140	150	160	170	180
	190	200	210	220	230	240
m121.pep	XXXXXXXXXXXXXX	(XXXXXXXXQI	PYDKNGAKS	AQGNILPQLL	RLLAHPYFA	ORHPKST
a121	:	1.1		:	11141111	LILLER
<b>a1</b> 21	PAFGFDTGPGNMLN	IDAMMQAHWQI	PYDKNGAKA	AQGNILPQLLD	RLLAHPYFA	QPHPKST
	190	200	210	220	230	240
	250	260	270	280	290	300
m121.pep	GRELFAINWLETYI	DGGENRYDVI	RTLSRFTAO	TVCDAVSHAAA	DAROMYTOD	יימואמדטר
	_	111111111			1111111	111111
a121	GRELFALNWLETYL	DGGENRYDVL	RTLSRFTAOT	VFDAVSHAAA	DAROMYTCG	CTDNDV
	250	260	270	280	290	300
	310	320	222			
m121.pep			330	340	350	360
mizi.pep	LMADLAECFGTRVS	LHSTADLNLD	POWVEAAXFA	WLAACWINRI	PGSPHKATGA	SKPCIL
a121				1:         :		11111
4424	LMADLAECFGTRVS 310	320	PQWVEAAAFA 330			
	310	320	330	340	350	360
m121.pep	XAGYYYX					
	11111					
a121	GAGYYYX	•				

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 441>: m121-1.seq

```
1 ATGGAAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
      GGCGGATGCC GTACTGATAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
  51
      AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTTACGCCG CCAATTGCTG
 101
 151
      GATTTGCAGG ACACAGGCGC AGACGAACTG CACCGCAGCA GGATTTTGTC
 201 GCAAGAACTC AGCCGCCTAT ATGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCACCGTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
 301 ACCGTCCGAC ACGCGCCGGA ACACGGTTAC AGCATACAGC TTGCCGATTT
 351
      GCCGCTGCTG GCGGAACGGA CGCGGATTTT TACCGTCGGC GACTTCCGCA
      GCCGCGACCT TGCGGCCGGC GGACAAGGCG CGCCACTCGT CCCCGCCTTT
 401
      CACGAAGCCC TGTTCCGCGA CAACAGGGAA ACACGCGCGG TACTGAACAT
 451
 501 CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCGACGCA CCCGCCTTCG
 551
      GCTTCGACAC AGGGCCGGGC AATATGCTGA TGGACGCGTG GACGCAGGCA
 601 CACTGGCAGC TTCCTTACGA CAAAAACGGT GCAAAGGCGG CACAAGGCAA
 651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCCACCCG TATTTCGCAC
 701 AACCCCACCC TAAAAGCACG GGGCGCGAAC TGTTTGCCCT AAATTGGCTC
751 GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TGCGGACGCT
 801
      TTCCCGTTTT ACCGCGCAAA CCGTTTGCGA CGCCGTCTCA CACGCAGCGG
      CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
 851
      TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
 901
      CACCGCCGAC CTGAACCTCG ATCCGCAATG GGTGGAAGCC GCCGNATTTG
 951
      CGTGGTTGGC GGCGTGTTGG ATTAATCGCA TTCCCGGTAG TCCGCACAAA
1051
      GCAACCGGCG CATCCAAACC GTGTATTCTG ANCGCGGGAT ATTATTATTG
1101
```

## This corresponds to the amino acid sequence <SEQ ID 442; ORF 121-1>: m121-1.pep

```
1 METQLYIGIM SGTSMDGADA VLIRMDGGKW LGAEGHAFTP YPGRLRRQLL
51 DLQDTGADEL HRSRILSQEL SRLYAQTAAE LLCSQNLAPS DITALGCHGQ
101 TVRHAPEHGY SIQLADLPLL AERTRIFTVG DFRSRDLAAG GQGAPLVPAF
151 HEALFRDNRE TRAVLNIGGI ANISVLPPDA PAFGFDTGFG NMLMDAWTQA
201 HWQLPYDKNG AKAAQGNLP QLLDRLLAHP YFAQPHFKST GRELFALNWL
251 ETYLDGGENR YDVLRTLSRF TAQTVCDAVS HAAADARQMY ICGGGIRNPV
301 LMADLAECFG TRVSLHSTAD LNLDPQWVEA AXFAWLAACW INRIPGSPHK
351 ATGASKPCIL XAGYYY*
```

m121-1/g121 95.6% identity in 366 aa overlap

10 20 30 40 50 60

m121-1.pep	METQLYIGIMSGTSMDGADAVLIRMDGGKWLGAEGHAFTPYPGRLRRQLLDLQD	
g121	METQLYIGIMSGTSMDGADAVLVRMDGGKWLGAEGHAFTPYPDRLRRKLLDLQD	
	70 80 90 100 110	120
m121-1.pep	HRSRILSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPEHGYSIQL	
q121		
9	70 80 90 100 110	120
	130 140 150 160 170	180
m121-1.pep	AERTRIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDNRETRAVLNIGGIANIS	
g121	AELTRIFTVGDFRSRDLAAGGOGAPLVPAFHEALFRDDRETRVVLNIGGIANIS	
9121	130 140 150 160 170	180
	190 200 210 220 230	240
m121-1.pep	PAFGFDTGPGNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQ	PHPKST
g121	PAFGFDTGPGNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLGRLLAHPYFSQ 190 200 210 220 230	240
	190 200 210 220 230	240
	250 260 270 280 290	300
m121-1.pep	GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVCDAVSHAAADARQMYICGG	
g121		
9121	250 260 270 280 290	300
	310 320 330 340 350	360
m121-1.pep	LMADLAECFGTRVSLHSTADLNLDPQWVEAAXFAWLAACWINRIPGSPHKATGA	
g121		
gizi	310 320 330 340 350	360
m121-1.pep	XAGYYYX	
• •	111111	
g121	GAGYYYX	

### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 443>: a121-1.seq

-		1				
	1	ATGGAAACAC	AGCTTTACAT	CGGCATCATG	TCGGGAACCA	GCATGGACGG
	51	GGCGGATGCC	GTACTGATAC	GGATGGACGG	CGGCAAATGG	CTGGGCGCGG
	101	AAGGGCACGC	CTTTACCCCC	TACCCCGGCA	GGTTACGCCG	CAAATTGCTG
	151	GATTTGCAGG	ACACAGGCGC	GGACGAACTG	CACCGCAGCA	GGATGTTGTC
	201	GCAAGAACTC	AGCCGCCTGT	ACGCGCAAAC	CGCCGCCGAA	CTGCTGTGCA
	251	GTCAAAACCT	CGCGCCGTCC	GACATTACCG	CCCTCGGCTG	CCACGGGCAA
	301	ACCGTCAGAC	ACGCGCCGGA	ACACAGTTAC	AGCGTACAGC	TTGCCGATTT
	351	GCCGCTGCTG	GCGGAACGGA	CTCAGATTTT	TACCGTCGGC	GACTTCCGCA
	401	GCCGCGACCT	TGCGGCCGGC	GGACAAGGCG	CGCCGCTCGT	CCCCGCCTTT
	451	CACGAAGCCC	TGTTCCGCGA	CGACAGGGAA	ACACGCGCGG	TACTGAACAT
	501	CGGCGGGATT	GCCAACATCA	GCGTACTCCC	CCCCGACGCA	CCCGCCTTCG
	551	GCTTCGACAC	AGGACCGGGC	AATATGCTGA	TGGACGCGTG	GATGCAGGCA
	601	CACTGGCAGC	TTCCTTACGA	CAAAAACGGT	GCAAAGGCGG	CACAAGGCAA
	651	CATATTGCCG	CAACTGCTCG	ACAGGCTGCT	CGCCCACCCG	TATTTCGCAC
	701	AACCCCACCC	TAAAAGCACG	GGGCGCGAAC	TGTTTGCCCT	AAATTGGCTC
	751	GAAACCTACC	TTGACGGCGG	CGAAAACCGA	TACGACGTAT	TGCGGACGCT
	801	TTCCCGATTC	ACCGCGCAAA	CCGTTTTCGA	CGCCGTCTCA	CACGCAGCGG
	851	CAGATGCCCG	TCAAATGTAC	ATTTGCGGCG	GCGGCATCCG	CAATCCTGTT
	901	TTAATGGCGG	ATTTGGCAGA	ATGTTTCGGC	ACACGCGTTT	CCCTGCACAG
	951	CACCGCCGAA	CTGAACCTCG	ATCCGCAATG	GGTAGAAGCC	GCCGCGTTCG
:	1001	CATGGATGGC	GGCGTGTTGG	GTCAACCGCA	TTCCCGGTAG	TCCGCACAAA
:	1051	GCAACCGGCG	CATCCAAACC	GTGTATTCTG	GGCGCGGGAT	ATTATTATTG
:	1101	A				

This corresponds to the amino acid sequence <SEQ ID 444; ORF 121-1.a>: a121-1.pep

- 1 METQLYIGIM SGTSMDGADA VLIRMDGGKW LGAEGHAFTP YPGRLRRKLL
  51 DLQDTGADEL HRSRMLSQEL SRLYAQTAAE LLCSQNLAPS DITALGCHGQ
  101 TVRHAPEHSY SVQLADLPLL AERTQIFTVG DFRSRDLAAG GQGAPLVPAF
  151 HEALFRDDRE TRAVLNIGGI ANISVLPPDA PAFGFDTGPG NMLMDAWMQA

```
HWQLPYDKNG AKAAQGNILP QLLDRLLAHP YFAQPHPKST GRELFALNWL
         ETYLDGGENR YDVLRTLSRF TAQTVFDAVS HAAADARQMY ICGGGIRNPV
     251
     301 LMADLAECFG TRVSLHSTAE LNLDPQWVEA AAFAWMAACW VNRIPGSPHK
     351 ATGASKPCIL GAGYYY*
ml21-1/a121-1 96.4% identity in 366 aa overlap
                          20
                                                  50
m121-1.pep
           METQLYIGIMSGTSMDGADAVLIRMDGGKWLGAEGHAFTPYPGRLRRQLLDLQDTGADEL
           METQLYIGIMSGTSMDGADAVLIRMDGGKWLGAEGHAFTPYPGRLRRKLLDLQDTGADEL
a121 - 1
                          20
                                  30
                                          40
                  70
                          80
                                  90
                                         100
                                                 110
m121-1.pep
           HRSRILSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPEHGYSIQLADLPLL
           a121-1
           HRSRMLSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPEHSYSVQLADLPLL
                         80
                                 90
                                                 110
                 130
                         140
                                150
                                         160
                                                 170
           AERTRIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDNRETRAVLNIGGIANISVLPPDA
m121-1.pep
           a121-1
           AERTQIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDDRETRAVLNIGGIANISVLPPDA
                 130
                        140
                                        160
                                                 170
                                210
                                        220
           PAFGFDTGPGNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST
m121-1.pep
           a121-1
           PAFGFDTGPGNMLMDAWMQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST
                190
                        200
                                210
                                        220
                250
                        260
                                270
                                        280
                                                290
                                                        300
           GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVCDAVSHAAADARQMYICGGGIRNPV
m121-1.pep
           a121-1
           GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVFDAVSHAAADARQMYICGGGIRNPV
                        260
                                270
                                        280
                310
                        320
                                330
                                        340
                                                350
m121-1.pep
          LMADLAECFGTRVSLHSTADLNLDPQWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL
           a121
          LMADLAECFGTRVSLHSTAELNLDPQWVEAAAFAWMAACWVNRIPGSPHKATGASKPCIL
                310
                        320
                                330
                                        340
                                                350
                                                        360
m121-1.pep
          XAGYYYX
           111111
a121
          GAGYYYX
```

### The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 445>: g122.seq

```
ATGGCTTTAC TGAGCATCCG CAAGCTGCAC AAACAATACG GCAGCGTAAC
     CGCCATCCAA TCCTTAGACT TGGACTTGGA AAAAGGCGAA GtcatCGTAC
 51
     TGCTGGGCCC gTccggctgc ggCAAATCCA CCCTcctgcg ctgcgtcaaC
101
     GGTTTGGAGC CGCACCAagg cgGCAGCATC GTGATGGACG GTgtcgGCGA
     ATTCggcAAA GACGTTTCCT GGCAAACCGC CCGGCAAAAa gtcggtatgg
251
     tctttcaaag taacgAactg Tttgcccaca tgaccgtcat cgAaaacatc
     ttcttAggcC CGGTAAagga aCAAAAcCgc gaccgtgccg aagcaGAGGC
301
351
     gCAAGCCGGC AAactGttgg aacgcgTCGG actgctAGAC CGCAAAAACG
401
     CCTATCCGCG CGAACTTTCC GGCGGTCAGA AACAGCGCAT CGCCATTGTC
     CGCGCCCTGT GCCTGAATCC GGAAGTCATC CTGCTGGACG AAATCACCGC
     CGCACTTGAC CCCGAAATGG TGCGCGAAGT CTTGGAAGTG GTTTTGGAAC
501
     TCGCCCGCGA AGGGATGAGT ATGCTCATCG TAACCCACGA AATGGGGTTC
551
     GCACGCAAAG TTGCCGACCG CATCGTCTTT ATGGACAAAG GCGGCATCGT
-601
     CGAATCGTCC GACCCCGAAA CCTTTTTTC CGCACCAAAA AGCGAACGCG
651
     CCCGCCAATT TCTGGCAGGT ATGGACTACT GA
```

This corresponds to the amino acid sequence <SEQ ID 446; ORF 122.ng>:

g122.pep

```
MALLSIRKLH KQYGSVTAIQ SLDLDLEKGE VIVLLGPSGC GKSTLLRCVN
           1
              GLEPHQGGSI VMDGVGEFGK DVSWQTARQK VGMVFQSNEL FAHMTVIENI
          51
             FLGPVKEQNR DRAEAEAQAG KLLERVGLLD RKNAYPRELS GGQKQRIAIV
         101
              RALCLNPEVI LLDEITAALD PEMVREVLEV VLELAREGMS MLIVTHEMGF
         201
              ARKVADRIVF MDKGGIVESS DPETFFSAPK SERARQFLAG MDY*
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 447>:
     m122.seq
           1
              GTTGTCATGA TTAAAATCCG CAATATCCAT AAGACCTTTG GCGAAAACAC
              TATTTTGCGC GGCATCGATT TGGATGTGTG CAAAGGGCAG GTGGTCGTCA
          51
              TCCTCGGGCC TTCCGGCTCA GGCAAAACGA CGTTTCTGCG ATGCCTAAAC
         101
         151 GCGTTGGAAA TGCCCGAAGA CGGACAAATC GAGTTCGACA ACGAGCGACC
              GCTGAAAATC GATTTTCTA AAAAACCAAG CAAACACGAT ATTTTGGCAC
             TGCGCCGCAA ATCAKGCATG GTGTTTCAAC AATACAAYCT CTTTCCGCAC
         251
         301 AAAACCGCCT TGGAAAACGT AATGGAAGGA CCGGTTGCCG TACAGGGCAA
         351 GCCTGCCGCC CAAGCGCGCG AAGAGGCTCT GAAACTGCTG GAAAAAGTCG
         401 GCTTGGGCGA CAAAGTGGAT TTGTATCCCT ACCAGCTTTC CGGCGGTCAG
         451 CAGCAGCGCG TCGGCATTGC CCGCGCATTG GCGATTCAGC CTGAACTGAT
         501 GCTGTTTGAC GAACCGACTT CCGCGCTCGA TCCTGAATTG GTGCAAGATG
         551 TTTTGGATmC CATGAAGGAA TTGGCGCAAG AAGGCTGGAC CATGGTTGTC
         601 GTTACGCATG AAATCAAGTT CGCCTTAGAA GTGGCAACCA CCGWCGTCGT
         651 GATGGACTGC GGCGTTATTG TCGAACAAGG CAGCCCGCAA GATTTGTTCG
         701 ACCACCCCAA ACACGAACGG ACGCGGAGAT TTTTAAGCCA AATCCAATCT
         751 ACCAAGATTT GA
This corresponds to the amino acid sequence <SEQ ID 448; ORF 122>:
     m122.pep
              VVMIKIRNIH KTFGENTILR GIDLDVCKGQ VVVILGPSGS GKTTFLRCLN
           1
          51 ALEMPEDGQI EFDNERPLKI DFSKKPSKHD ILALRRKSXM VFQQYNLFPH
             KTALENVMEG PVAVQGKPAA QAREEALKLL EKVGLGDKVD LYPYQLSGGQ
         151 OQRVGIARAL AIQPELMLFD EPTSALDPEL VQDVLDXMKE LAQEGWTMVV
         201 VTHEIKFALE VATTXVVMDX GVIVEQGSPQ DLFDHPKHER TRRFLSQIQS
         251
              TKI *
Computer analysis of this amino acid sequence gave the following results:
Homology with a predicted ORF from N. gonorrhoeae
ORF 122 shows 47.2% identity over a 246 aa overlap with a predicted ORF (ORF 122.ng)
from N. gonorrhoeae:
     m122/q122
                                  20
                                           30
                                                    40
                 VVMIKIRNIHKTFGENTILRGIDLDVCKGQVVVILGPSGSGKTTFLRCLNALEMPEDGQI
     m122.pep
                 : |:|
                 MALLSIRKLHKQYGSVTAIQSLDLDLEKGEVIVLLGPSGCGKSTLLRCVNGLEPHQGGSI
     q122
                        10
                                  20
                                           30
                                                    40
                                                              50
                                                                       60
                        70
                                  80
                                           90
                                                   100
                 EFDNERPLKIDFSKKPSKHDILALRRKSXMVFQQYNLFPHKTALENVMEGPVAVQGKPAA
     m122.pep
                       : | | : :
                                      VMDGVGEFGKDVSWQTA-----RQKVGMVFQSNELFAHMTVIENIFLGPVKEQNRDRA
     g122
                        70
                                                 90
                                                          100
                                                                   110
                                          150
                                                   160
                                                             170
                 QAREEALKLLEKVGLGDKVDLYPYQLSGGQQQRVGIARALAIQPELMLFDEPTSALDPEL
     m122.pep
                 EAEAQAGKLLERVGLLDRKNAYPRELSGGQKQRIAIVRALCLNPEVILLDEITAALDPEM
     q122
                     120
                              130
                                       140
                                                150
                                                          160
                                                                   170
                                 200
                                          210
                                                   220
                                                             230
                 VQDVLDXMKELAQEGWTMVVVTHEIKFALEVATTXVVMDXGVIVEQGSPQDLFDHPKHER
     m122.pep
                 g122
                 VREVLEVVLELAREGMSMLIVTHEMGFARKVADRIVFMDKGGIVESSDPETFFSAPKSER
```

BNSDOCID: <WO___9957280A2_I_>

180 190 200 210 220 230 250 m122.pep TRRFLSQIQSTKIX : | : | | : q122 ARQFLAGMDYX 240 The following partial DNA sequence was identified in N. meningitidis <SEQ ID 449>: a122.seg GTTGTCATGA TTAAAATCCG CAATATCCAT AAGACCTTCG GCAAAAATAC 51 CATTTTGCGC GGCATCAATT TGGATGTGTG CAAAGGGCAG GTGGTCGTCA TCCTCGGGCC TTCCGGCTCA GGCAAAACGA CGTTTCTGCG ATGCCTAAAC 101 151 GCGTTGGAAA TGCCCGAAGA CGGACAAATC GAGTTCGACA ACGAGCGACC GCTGAAAATC GATTTTCTA AAAAACCAAG CAAACACGAT ATTTTGGCAC 201 TGCGCCGCAA ATCAGGCATG GTGTTTCAAC AATACAACCT CTTTCCGCAC 251 AAAACCGCCT TGGAAAACGT GATGGAAGGA CCGGTTGCCG TACAGGGCAA 301 GCCTGCCGCC CAAGCGCGCG AAGAGGCTCT GAAACTGCTG GAAAAAGTCG 351 401 GCTTGGGCGA CAAAGTGGAT TTGTATCCCT ACCAGCTTTC CGGCGGTCAG 451 CAGCAGCGCG TCGGCATTGC CCGAGCATTG GCGATTCAGC CCGAGCTGAT GTTGTTTGAC GAACCCACTT CCGCGCTTGA CCCCGAGTTG GTGCAAGACG 501 TGTTGAACGC CATGAAGGAA TTGGCGCGGG AAGGTTGGAC GATGGTCGTC 551 GTTACCCACG AAATCAAGTT CGCGCTGGAA GTTGCCACGA CCGTTGTCGT 601 GATGGACGGC GGCGTTATCG TAGAGCAGGG CAGCCCGAAA GAGTTGTTCG 651 ACCACCCCAA ACACGAACGG ACGCGGAGAT TTTTAAGCCA AATCCAATCT 701 751 ACCAAGATTT GA This corresponds to the amino acid sequence <SEQ ID 450; ORF 122.a>: al22.pep VVMIKIRNIH KTFGKNTILR GINLDVCKGQ VVVILGPSGS GKTTFLRCLN ALEMPEDGQI EFDNERPLKI DFSKKPSKHD ILALRRKSGM VFQQYNLFPH 51 101 KTALENVMEG PVAVQGKPAA QAREEALKLL EKVGLGDKVD LYPYQLSGGO 151 QQRVGIARAL AIQPELMLFD EPTSALDPEL VQDVLNAMKE LAREGWTMVV VTHEIKFALE VATTVVVMDG GVIVEQGSPK ELFDHPKHER TRRFLSQIQS 201 251 TKI* m122/a122 96.0% identity in 253 aa overlap 10 20 30 40 VVMIKIRNIHKTFGENTILRGIDLDVCKGQVVVILGPSGSGKTTFLRCLNALEMPEDGQI m122.pep VVMIKIRNIHKTFGKNTILRGINLDVCKGQVVVILGPSGSGKTTFLRCLNALEMPEDGQI a122 10 20 30 40 50 60 80 90 100 110 EFDNERPLKIDFSKKPSKHDILALRRKSXMVFQQYNLFPHKTALENVMEGPVAVQGKPAA m122.pep EFDNERPLKIDFSKKPSKHDILALRRKSGMVFQQYNLFPHKTALENVMEGPVAVQGKPAA a122 70 80 90 100 110 120 130 140 150 160 170 QAREEALKLLEKVGLGDKVDLYPYQLSGGQQQRVGIARALAIQPELMLFDEPTSALDPEL m122.pep QAREEALKLLEKVGLGDKVDLYPYQLSGGQQQRVGIARALAIQPELMLFDEPTSALDPEL a122 130 140 150 160 190 200 210 220 230 VQDVLDXMKELAQEGWTMVVVTHEIKFALEVATTXVVMDXGVIVEQGSPQDLFDHPKHER m122.pep VQDVLNAMKELAREGWTMVVVTHEIKFALEVATTVVVMDGGVIVEQGSPKELFDHPKHER a122 190 200 210 220 230 250 m122.pep TRRFLSQIQSTKIX

al22 TRRFLSQIQSTKIX 250

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 451>: g122-1.seq

```
1 ATGATTAAAA TCCGCAATAT CCATAAGACC TTTGGCGAAA ACACCATTTT
     GCGCGGCATC GATTTGGATG TGGGCAAAGG GCAGGTGGTC GTCATCCTCG
 51
101 GGCCTTCCGG CTCGGGTAAA ACAACATTTC TGCGCTGCCT AAACGCGTTG
151 GAAATGCCCG AAGACGGACA AATCGAGTTC GACAACGCGC GGCCGTTACG
201 CATTGATTTT TCCAAAAAAA CAAGCAAACA CGATATTTTG GCACTGCGCC
251 GCAAGTCCGG AATGGTATTC CAACAATACA ACCTCTTCCC GCATAAAACC
301 GTGTTGGAAA ACGTGATGGA AGGGCCGGTT GCCGTACAGG GCAAGCCTGC
351 CGCCCAAGCG CGCGAAGAGG CTTTGAAACT GCTGGAAAAA GTCGGCTTGG
    GCGATAAAGT GGATTTGTAT CCCTACCAGC TTTCCGGCGG TCAGCAGCAG
CGTGTCGGTA TCGCCCGCGC ACTGGCGATT CAGCCTGAAT TGATGCTGTT
401
451
501 TGACGAACCC ACTTCCGCGC TGGACCCCGA GTTGGTGCAA GACGTGTTGG
    ACGCCATGAA GGAATTGGCG CGGGAAGGTT GGACGATGGT CGTCGTTACC
601 CACGAAATCA AGTTCACGCT GGAAGTTGCC ACGAACGTCG TCGTGATGGA
    CGGCGGCGTT ATCGTAGAGC AGGGCAGCCC GAAAGAGTTG TTCGACCACC
701 TCAAACACGA ACGGACGCGG AGATTTTTAA GCCAAATCCA ATCTGCCAAG
751 ATTTGA
```

This corresponds to the amino acid sequence <SEQ ID 452; ORF 122-1.ng>: g122-1.pep

```
1 MIKIRNIHKT FGENTILRGI DLDVGKGQVV VILGPSGSGK TTFLRCLNAL
51 EMPEDGQIEF DNARPLRIDF SKKTSKHDIL ALRRKSGMVF QQYNLFPHKT
101 VLENVMEGPV AVQGKPAAQA REEALKLLEK VGLGDKVDLY PYQLSGGQQQ
151 RVGIARALAI QPELMLFDEP TSALDPELVQ DVLDAMKELA REGWTMVVVT
201 HEIKFTLEVA TNVVVMDGGV IVEQGSPKEL FDHLKHERTR RFLSQIQSAK
```

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 453>: m122-1.seq

```
1 ATGATTAAAA TCCGCAATAT CCATAAGACC TTTGGCGAAA ACACTATTTT
 51 GCGCGGCATC GATTTGGATG TGTGCAAAGG GCAGGTGGTC GTCATCCTCG
101 GGCCTTCCGG CTCAGGCAAA ACGACGTTTC TGCGATGCCT AAACGCGTTG
151 GAAATGCCCG AAGACGGACA AATCGAGTTC GACAACGAGC GACCGCTGAA
201 AATCGATTTT TCTAAAAAAC CAAGCAAACA CGATATTTTG GCACTGCGCC
    GCAAATCAGG CATGGTGTTT CAACAATACA ACCTCTTTCC GCACAAAACC
301 GCCTTGGAAA ACGTAATGGA AGGACCGGTT GCCGTACAGG GCAAGCCTGC
    CGCCCAAGCG CGCGAAGAGG CTCTGAAACT GCTGGAAAAA GTCGGCTTGG
351
401 GCGACAAAGT GGATTTGTAT CCCTACCAGC TTTCCGGCGG TCAGCAGCAG
451 CGCGTCGGCA TTGCCCGCGC ATTGGCGATT CAGCCTGAAC TGATGCTGTT
    TGACGAACCG ACTTCCGCGC TCGATCCTGA ATTGGTGCAA GATGTTTTGG
551 ATACCATGAA GGAATTGGCG CAAGAAGGCT GGACCATGGT TGTCGTTACG
    CATGAAATCA AGTTCGCCTT AGAAGTGGCA ACCACCGTCG TCGTGATGGA
651 CGGCGGCGTT ATTGTCGAAC AAGGCAGCCC GCAAGATTTG TTCGACCACC
701
    CCAAACACGA ACGGACGCGG AGATTTTTAA GCCAAATCCA ATCTACCAAG
    ATTTGA
```

This corresponds to the amino acid sequence <SEQ ID 454; ORF 122-1>: m122-1.pep

```
1 MIKIRNIHKT FGENTILRGI DLDVCKGQVV VILGPSGSGK TTFLRCLNAL
51 EMPEDGQIEF DNERPLKIDF SKKPSKHDIL ALRRKSGMVF QQYNLFPHKT
101 ALENVMEGPV AVQGKPAAQA REEALKLLEK VGLGDKVDLY PYQLSGGQQQ
151 RGIARALAI QPELMLFDEP TSALDPELVQ DVLDTMKELA QEGWTMVVVT
201 HEIKFALEVA TTVVVMDGGV IVEQGSPQDL FDHPKHERTR RFLSQIQSTK
```

m122-1/g122-1 94.8% identity in 251 aa overlap

BNSDOCID: <₩O__9957280A2_]>

	10	20	30	40	50	60
m122-1.pep	70 DNERPLKIDFSKI	80 PSKHDILALRI	90 RKSGMVFQQYN	100 ILFPHKTALEN	110 VVMEGPVAVQO	120 GKPAAQA
g122-1	DNARPLRIDFSKE 70	 TSKHDILALRI 80	RKSGMVFQQYN 90	:     LFPHKTVLEN   100	IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	  KPAAQA   120
m122-1.pep	130 REEALKLLEKVGL            REEALKLLEKVGL 130	1111111111	111111111	311111111	 MLFDEPTSAI	DPELVQ
m122-1.pep g122-1	190 DVLDTMKELAQEG	200 WTMVVVTHEIK	210 FALEVATTVV	220 VMDGGVIVEQ	111::111	111111
m122-1.pep	250 RFLSQIQSTKIX          RFLSQIQSAKIX 250				230	2. <del>1</del> U
The C-11				. ~		

### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 455>: a122-1.seq

```
1 ATGATTAAAA TCCGCAATAT CCATAAGACC TTCGGCAAAA ATACCATTTT 51 GCGCGGCATC AATTTGGATG TGTGCAAAGG GCAGGTGGTC GTCATCCTCG
101 GGCCTTCCGG CTCAGGCAAA ACGACGTTTC TGCGATGCCT AAACGCGTTG
151 GAAATGCCCG AAGACGGACA AATCGAGTTC GACAACGAGC GACCGCTGAA
201 AATCGATTTT TCTAAAAAAC CAAGCAAACA CGATATTTTG GCACTGCGCC
251 GCAAATCAGG CATGGTGTTT CAACAATACA ACCTCTTTCC GCACAAAACC 301 GCCTTGGAAA ACGTGATGGA AGGACCGGTT GCCGTACAGG GCAAGCCTGC
351 CGCCCAAGCG CGCGAAGAGG CTCTGAAACT GCTGGAAAAA GTCGGCTTGG
401 GCGACAAAGT GGATTTGTAT CCCTACCAGC TTTCCGGCGG TCAGCAGCAG
451 CGCGTCGGCA TTGCCCGAGC ATTGCCCGAGC TGATGTTGTT
501 TGACGAACCC ACTTCCGCGC TTGACCCCGA GTTGGTGCAA GACGTGTTGA
551 ACGCCATGAA GGAATTGGCG CGGGAAGGTT GGACGATGGT CGTCGTTACC
601 CACGAAATCA AGTTCGCGCT GGAAGTTGCC ACGACCGTTG TCGTGATGGA
651 CGGCGGCGTT ATCGTAGAGC AGGGCAGCCC GAAAGAGTTG TTCGACCACC
701 CCAAACACGA ACGGACGCGG AGATTTTTAA GCCAAATCCA ATCTACCAAG
751 ATTTGA
```

### This corresponds to the amino acid sequence <SEQ ID 456; ORF 122-1.a>: a122-1.pep

- 1 MIKIRNIHKT FGKNTILRGI NLDVCKGQVV VILGPSGSGK TTFLRCLNAL 51 EMPEDGQIEF DNERPLKIDF SKKPSKHDIL ALRRKSGMVF QQYNLFPHKT 101 ALENVMEGPV AVQGKPAAQA REEALKLLEK VGLGDKVDLY PYQLSGGQQQ 151 RVGIARALAI QPELMLFDEP TSALDPELVQ DVLNAMKELA REGWTMVVVT 201 HEIKFALEVA TTVVVMDGGV IVEQGSPKEL FDHPKHERTR RFLSQIQSTK
- a122-1/m122-1 97.2% identity in 251 aa overlap

251 I*

a122-1.pep	10 MIKIRNIHKTFGKN             MIKIRNIHKTFGEN	111111:111		1111111111		
	10	20	CKGOAATTO			
	10	20	30	40	50	60
a122-1.pep	70 DNERPLKIDFSKKP             DNERPLKIDFSKKP 70	111111111		1111111111	1111111111	11111
a122-1.pep	130 REEALKLLEKVGLG !!!!!!!!!!! REEALKLLEKVGLG		11111111	11111111111	1111111111	11111

	130	140	150	160	170	180
a122-1.pep m122-1	190 DVLNAMKELAREGW    ::    ::    DVLDTMKELAQEGW 190	111111111111111111111111111111111111111		11111111		1111111
al22-1.pep	250 RFLSQIQSTKIX            RFLSQIQSTKIX 250					

### The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 457>: g125.seq

```
ATGTCGGGCA ATGCCTCCTC TCCTTCATCT TCCGCCGCCA TCGGGCTGGT
  1
 53
     TTGGTTCGGC GCGCGGTAT CGATTGCCGA AATCAGCACG GGTACGCTGC
101 TCGCCCCTT GGGCTGGCAG CGCGGTCTGG CGCCCTGCT TTTGGGTCAT
151 GCCGTCGGCG GCGCGCTGTT TTTTGCGGCG GCGTATATCG GCGCACTGAC
201 CGGACGCAGC TCGATGGAAA GTGTGCGCCT GTCGTTCGGC AAATGCGGTT
251 CAGTGCTGTT TTCCGTGGCG AATATGCTGC AACTGGCCGG CTGGACGGCG
301 GTGATGATTT ACGTCGGCGC AacggTCAGC TCCGCTTTGG GCAAAGTGTT
351 GTGGGACggc gaATCCTTTG TCTGGTGGGC ATTGGCAAAC GGCGCACTGA
401 TCGTGCTGTG GCTGGTTTTC GGCGCACGCA GAACGGGCGG GCTGAAAACC
451 GTTTCGATGC TGCTGATGCT GCTTGCCGTG TTGTGGTTGA GCGTCGAAGT
     GTTCGCTTCG TCCGGCACAA ACGCCGCGCC CGCCGTTTCA GACGGCATGA
    CCTTCGGAAC GGCAGTCGAA CTGTCCGCCG TCATGCCGCT TTCCTGGCTG
    CCGCTGGCCG CCGACTACAC GCGCCAAGCA CGCCGCCCGT TTGCGGCAAC
651 CCTGACGGCA ACGCTCGCCT ATACGCTGAC GGGCTGCTGG ATGTATGCCT
701 TGGGTTTGGC GGCGGCTCTG TTTACCGGAG AAACCGACGT GGCGAAAATC
751 CTGTTGGGCG CGGGCTTGGG CATAACGGGC ATTCTGGCAG TCGTCCTCTC
801 CACCGTTACC ACAACGTTTC TCGATACCTA TTCCGCCGGC GCGAGTGCGA
851 ACAACATTC CGCGCGTTTT GCGGAAATAC CCGTCGCTGT CGGCGTTACC
901 CTGatccgca ccgtgcttgc cgtcatgctg cccgttaccg aatataaaaa
951 cttcctgctg cttatccgct cggtatttgg gccgatggcg ggtggttttg
1001 attgccgaCT TTTttgtctt AAAACGGCGT GA
```

This corresponds to the amino acid sequence <SEQ ID 458; ORF 125.ng>:

#### The following partial DNA sequence was identified in N. meningitidis <SEO ID 459>:

```
m125.sea
         ATGTCGGGCA ATGCCTCCTC TCCTTCATCT TCCTCCGCCA TCGGGCTGAT
      1
     51
         TTGGTTCGGC GCGGCGGTAT CGATTGCCGA AATCAGCACG GGTACGCTGC
    101 TTGCGCCTTT GGGCTGGCAG CGCGGTCTGG CGGCTCTACT TTTGGGTCAT
    151 GCCGTCGGCG GCGCGCTGTT TTTTGCGGCG GCGTATATCG GCGCACTGAC
         CGGACGCAGC TCGATGGAAA GCGTGCGCCT GTCGTTCGGC AAACGCGGTT
         CAGTGCTGTT TTCCGTGGCG AATATGCTGC AACTGGCCGG CTGGACGGCG
         GTGATGATTT ACGCCGCGC AACGGTCAGC TCCGCTTTGG GCAAAGTGTT
         GTGGGACGCC GAATCTTTTG TCTGGTGGGC ATTGGCAAAC GGCGCGCTGA
    401 TTGTGCTGTG GCTGGTTTTC GGCGCACGCA AAACAGGCGG GCTGAAAACC
    451 GTTTCGATGC TGCTGATGCT GTTGGCGGTT CTGTGGCTGA GTGCCGAAGT
    501 CTTTTCCACG GCAGGCAGCA CCGCCGCACA GGTTTCAGAC GGCATGAGTT
    551 TCGGAACGGC AGTCGAGCTG TCCGCCGTGA TGCCGCTTTC CTGGCTGCCG
    601 CTTGCCGCCG ACTACACGCG CCACGCGCGC CGCCCGTTTG CGGCAACCCT
    651 GACGGCAACG CTCGCCTACA CGCTGACCGG CTGCTGGATG TATGCCTTGG
    701 GTTTGGCAGC GGCGTTGTTC ACCGGAGAAA CCGACGTGGC AAAAATCCTG
    751 CTGGGCGCAr GTTTGGGTGC GGCAGGCATT TTGGCGGTCG TCCTCTCCAC
    801 CGTTACCACA ACGTTTCTCG ATGCCTATTC CGCCGGCGCG AGTGCGAACA
```

```
851 ACATTTCCGC GCGTTTTGCG GAAACACCCG TCGCTGTCrG CGTTACCCTG
             ATCGGCACGG TACTTGCCGT CATGCTGCCC GTTACCGAAT ATGAAAACTT
              CCTGCTGCTT ATCGGCTCGG TATTTGCGCC GATGGCGGGC GGTTTTGATT
              GCCGACTTTT TCGTCTTGAA ACGGCGTGA
This corresponds to the amino acid sequence <SEQ ID 460; ORF 125>:
     m125.pep
             MSGNASSPSS SSAIGLIWFG AAVSIAEIST GTLLAPLGWQ RGLAALLLGH
             AVGGALFFAA AYIGALTGRS SMESVRLSFG KRGSVLFSVA NMLQLAGWTA
          51
             VMIYAGATVS SALGKVLWDG ESFVWWALAN GALIVLWLVF GARKTGGLKT
         101
             VSMLLMLLAV LWLSAEVFST AGSTAAQVSD GMSFGTAVEL SAVMPLSWLP
             LAADYTRHAR RPFAATLTAT LAYTLTGCWM YALGLAAALF TGETDVAKIL
             LGAXLGAAGI LAVVLSTVTT TFLDAYSAGA SANNISARFA ETPVAVXVTL
         251
             IGTVLAVMLP VTEYENFLLL IGSVFAPMAG GFDCRLFRLE TA*
         301
Computer analysis of this amino acid sequence gave the following results:
Homology with a predicted ORF from N. gonorrhoeae
ORF 125 shows 92.1% identity over a 343 aa overlap with a predicted ORF (ORF 125.ng)
from N. gonorrhoeae:
    m125/g125
                                        30
                                                 40
                                                          50
               MSGNASSPSSSSAIGLIWFGAAVSIAEISTGTLLAPLGWQRGLAALLLGHAVGGALFFAA
    m125.pep
                MSGNASSPSSSAAIGLVWFGAAVSIAEISTGTLLAPLGWQRGLAALLLGHAVGGALFFAA
    q125
                               20
                                        30
                                                 40
                                                          50
                       70
                               80
                                        90
                                                100
                                                         110
    m125.pep
               AYIGALTGRSSMESVRLSFGKRGSVLFSVANMLQLAGWTAVMIYAGATVSSALGKVLWDG
               AYIGALTGRSSMESVRLSFGKCGSVLFSVANMLQLAGWTAVMIYVGATVSSALGKVLWDG
    g125
                      70
                               80
                                        90
                                                100
                                                         110
                     130
                              140
                                       150
                                                         170
                                                                  179
               ESFVWWALANGALIVLWLVFGARKTGGLKTVSMLLMLLAVLWLSAEVFSTAGSTAAQ-VS
    m125.pep
               ESFVWWALANGALIVLWLVFGARRTGGLKTVSMLLMLLAVLWLSVEVFASSGTNAAPAVS
   g125
                     130
                              140
                                       150
                                                160
                                                         170
             180
                      190
                               200
                                        210
                                                 220
                                                         230
                                                                 239
               DGMSFGTAVELSAVMPLSWLPLAADYTRHARRPFAATLTATLAYTLTGCWMYALGLAAAL
   m125.pep
               DGMTFGTAVELSAVMPLSWLPLAADYTROARRPFAATLTATLAYTLTGCWMYALGLAAAL
   g125
                     190
                              200
                                       210
                                                220
                                                        230
                                                                 240
             240
                      250
                               260
                                        270
                                                280
                                                         290
                                                                 299
               FTGETDVAKILLGAXLGAAGILAVVLSTVTTTFLDAYSAGASANNISARFAETPVAVXVT
   m125.pep
               FTGETDVAKILLGAGLGITGILAVVLSTVTTTFLDTYSAGASANNISARFAEIPVAVGVT
   g125
                     250
                              260
                                      270
                                               280
                                                        290
                                                                 300
```

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 461>:

320

320

330

330

340

300

m125.pep

g125

310

310

1 ATGTCGGGCA ATGCCTCCTC TCCTTCATCT TCCGCCGCCA TCGGGCTGAT
51 TTGGTTCGGC GCGGCGGTAT CGATTGCCGA AATCAGCACG GGTACACTGC

101 TTGCGCCTTT GGGCTGGCAG CGCGGTCTGG CGGCTCTGCT TTTGGGTCAT

151	GCCGTCGGCG GCGCGCTGTT TTTTGCGGCG GCGTATATCG GCGCACTGAC	
201	CGGACGCAGC TCGATGGAAA GCGTGCGCCT GTCGTTCGGC AAACGCGGTT	
251	CAGTGCTGTT TTCCGTGGCG AATATGCTGC AACTGGCCGG CTGGACGGCG	
301	GTGATGATTT ACGCCGGCGC AACGGTCAGC TCCGCTTTGG GCAAAGTGTT	
351	GTGGGACGGC GAATCTTTTG TCTGGTGGGC ATTGGCAAAC GGCGCGCTGA	
401	TTGTGCTGTG GCTGGTTTTC GGCGCACGCA AAACAGGCGG GCTGAAAACC	
451	GTTTCGATGC TGCTGATGCT GTTGGCGGTT CTGTGGCTGA GTGCCGAAGT	
501	CTTTTCCACG GCAGGCAGCA CCGCCGCACA GGTTTCAGAC GGCATGAGTT	
551	TCGGAACGGC AGTCGAGCTG TCCGCCGTGA TGCCGCTTTC TTGGCTGCCG	
601	CTGGCCGCCG ACTACACGCG CCACGCGCGC CGCCCGTTTG CGGCAACCCT	
651	GACGGCAACG CTCGCCTACA CGCTGACCGG CTGCTGGATG TATGCCTTGG	
701	GTTTGGCAGC GGCGTTGTTC ACCGGAGAAA CCGACGTGGC AAAAATCCTG	
751	CTGGGCGCAG GTTTGGGTGC GGCAGGCATT TTGGCGGTCG TCCTGTCGAC	
801	CGTTACCACC ACTTTTCTCG ATGCCTACTC CGCCGGCGTA AGTGCCAACA	
851	ATATTTCCGC CAAACTTTCG GAAATACCCA TCGCCGTTGC CGTCGCCGTT	
901	GTCGGCACAC TGCTTGCCGT CCTCCTGCCC GTTACCGAAT ATGAAAACTT	
951	CCTGCTGCTT ATCGGCTCGG TATTTGCGCC GATGGCG.GC GGTTTTGATT	
1001	GCCGACTTTT TCGTCTTGAA ACGGCGTGA	
This correspond	s to the amino acid sequence <seq 125.a="" 462;="" id="" orf="">:</seq>	
al25.pep	to the attimo acta dequation of the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first in the first index in the first in the first in the first in the first in the fi	
a125.pep	MSGNASSPSS SAAIGLIWFG AAVSIAEIST GTLLAPLGWO RGLAALLLGH	
51	AVGGALFFAA AYIGALTGRS SMESVRLSFG KRGSVLFSVA NMLQLAGWTA	
101	VMIYAGATVS SALGKVLWDG ESFVWWALAN GALIVLWLVF GARKTGGLKT	
151	VSMLIMLLAV LWLSAEVFST AGSTAAQVSD GMSFGTAVEL SAVMPLSWLP	
201	LAADYTRHAR RPFAATLTAT LAYTLTGCWM YALGLAAALF TGETDVAKIL	
251		
	LGAGLGAAGI LAVVLSTVTT TFLDAYSAGV SANNISAKLS EIPIAVAVAV	
301	VGTLLAVLLP VTEYENFLLL IGSVFAPMAX GFDCRLFRLE TA*	
125/-125 06	COV identify in 242 or examina	
m125/a125 95	6.6% identity in 342 aa overlap	
	10 20 30 40 50	60
m125/a125 95	10 20 30 40 50 MSGNASSPSSSSAIGLIWFGAAVSIAEISTGTLLAPLGWQRGLAALLLGHAVGGALF	AA
m125.pep	10 20 30 40 50 MSGNASSPSSSSAIGLIWFGAAVSIAEISTGTLLAPLGWQRGLAALLLGHAVGGALF	FAA 
	10 20 30 40 50  MSGNASSPSSSAIGLIWFGAAVSIAEISTGTLLAPLGWORGLAALLLGHAVGGALFI	FAA 
m125.pep	10 20 30 40 50 MSGNASSPSSSSAIGLIWFGAAVSIAEISTGTLLAPLGWQRGLAALLLGHAVGGALF	FAA 
m125.pep	10 20 30 40 50  MSGNASSPSSSAIGLIWFGAAVSIAEISTGTLLAPLGWORGLAALLLGHAVGGALFI	FAA       FAA
m125.pep a125	10 20 30 40 50  MSGNASSPSSSAIGLIWFGAAVSIAEISTGTLLAPLGWORGLAALLLGHAVGGALF!	FAA 111 FAA 60
m125.pep	10 20 30 40 50  MSGNASSPSSSAIGLIWFGAAVSIAEISTGTLLAPLGWORGLAALLLGHAVGGALFI	FAA 60 L20 VDG
m125.pep a125 m125.pep	10 20 30 40 50  MSGNASSPSSSAIGLIWFGAAVSIAEISTGTLLAPLGWORGLAALLLGHAVGGALF!	FAA 60 L20 VDG
m125.pep a125	10 20 30 40 50  MSGNASSPSSSAIGLIWFGAAVSIAEISTGTLLAPLGWORGLAALLLGHAVGGALF;	FAA 60 L20 VDG
m125.pep a125 m125.pep	10 20 30 40 50  MSGNASSPSSSAIGLIWFGAAVSIAEISTGTLLAPLGWORGLAALLLGHAVGGALFI	FAA 60 L20 VDG
m125.pep a125 m125.pep	10 20 30 40 50  MSGNASSPSSSAIGLIWFGAAVSIAEISTGTLLAPLGWQRGLAALLLGHAVGGALF;	FAA 60 L20 VDG III VDG
m125.pep a125 m125.pep a125	10 20 30 40 50  MSGNASSPSSSAIGLIWFGAAVSIAEISTGTLLAPLGWQRGLAALLLGHAVGGALF!	FAA 111 FAA 60 L20 VDG 111 VDG L20
m125.pep a125 m125.pep	10 20 30 40 50  MSGNASSPSSSAIGLIWFGAAVSIAEISTGTLLAPLGWORGLAALLLGHAVGGALFI	FAA 60 L20 NDG III NDG L20
m125.pep a125 m125.pep a125	10 20 30 40 50  MSGNASSPSSSAIGLIWFGAAVSIAEISTGTLLAPLGWORGLAALLLGHAVGGALFI	FAA 60 L20 NDG III NDG L20 L80
m125.pep a125 m125.pep a125	10 20 30 40 50  MSGNASSPSSSAIGLIWFGAAVSIAEISTGTLLAPLGWORGLAALLLGHAVGGALFI	FAA 1111 FAA 60 L20 VDG 1111 VDG L20 L80 VSD
m125.pep a125 m125.pep a125	10 20 30 40 50  MSGNASSPSSSAIGLIWFGAAVSIAEISTGTLLAPLGWORGLAALLLGHAVGGALFI	FAA 60 L20 NDG III NDG L20 L80
m125.pep a125 m125.pep a125	10 20 30 40 50  MSGNASSPSSSAIGLIWFGAAVSIAEISTGTLLAPLGWQRGLAALLLGHAVGGALFI	FAA 60 L20 WDG III WDG L20 L80 /SD III
m125.pep a125 m125.pep a125 m125.pep a125	10 20 30 40 50  MSGNASSPSSSAIGLIWFGAAVSIAEISTGTLLAPLGWQRGLAALLLGHAVGGALFI	FAA           FAA   60   L20   NDG   1       NDG   L20   L80   / SD           / SD   L80
m125.pep a125 m125.pep a125	10 20 30 40 50  MSGNASSPSSSAIGLIWFGAAVSIAEISTGTLLAPLGWQRGLAALLLGHAVGGALFI	FAA             FAA   60   120   120   100   100
m125.pep a125 m125.pep a125 m125.pep a125	10 20 30 40 50  MSGNASSPSSSAIGLIWFGAAVSIAEISTGTLLAPLGWQRGLAALLLGHAVGGALFI	FAA 111   FAA 60   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120
m125.pep a125 m125.pep a125 m125.pep a125	10 20 30 40 50  MSGNASSPSSSAIGLIWFGAAVSIAEISTGTLLAPLGWQRGLAALLLGHAVGGALFI	FAA 111   FAA 60   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120
m125.pep a125 m125.pep a125 m125.pep a125	10	FAA 111   FAA 60   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120
m125.pep a125 m125.pep a125 m125.pep a125	10	FAA 111   FAA 60   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120   120
m125.pep a125 m125.pep a125 m125.pep a125	10	FAA         FFAA 60 120 120 180 175D 180 180 180 180 180 180 180 180
m125.pep a125 m125.pep a125 m125.pep a125	10	FAA           FAA   60   120   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100
m125.pep a125 m125.pep a125 m125.pep a125	10	FAA           FAA   60   120   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100   100

TGETDVAKILLGAGLGAAGILAVVLSTVTTTFLDAYSAGVSANNISAKLSEIPIAVAVAV

330

270 280

340

290

BNSDOCID: <WO___9957280A2_i_>

a125

a125

m125.pep

250

310

260

320

320

330

340

# The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 463>:

```
1
     AtgccgtcTG AAaccCcaaa ggcACGCCGC CGGCTTTCAG ACGGCATCGC
     GTCCGACAAC CATACCAAAG AATCCATCAT GCTCACCCtg tacggcGAAA
 51
     CTTTCCCTTC GCGGCTGCTg ctcggcacgG cggcctacCC GACCCCTGAA
101
151 ATCCTCAAAC AATCCGTCCG AACCGCCCGG CCCGCGATGA ttaccGTCTC
     GCTGCGCCGC ACGGGATGCG GCGCGAGGC GCACGGTCAG GGGTTTTGGT
201
251
     CGCTGCTTCA AGAAACCGGC GTTCCCGTCC TGCCGAACAC GGCAGGCTGC
     CAAAGCGTGC AGGAAGCGGT. AACGACGGCG CAAATGGCGC GCGAAGTGTT
351 TGAAACCGAT TGGATAAAAT TGGAACTCAT CGGCGACGAC GACACCTTGC
401 AGCCGGACGT GTTCCAACTC GTCGAAGCGG CGGAAATCCT GATTAAAGAC
    GGCTTCAAAG TGCTGCCTTA TTGCACCGAA GACCTGATTG CCTGCCGCCG
451
     CCTGCTCGAT GCGGGCTGTC AGGCGTTGAT GCCGTGGGCG GCTCCCATCG
     GCACGGGTTT GGGGGCGGTT CACGCCTATG CGCTCAAAAT CCTGCGCGAA
     CGCCTGCCCG ACACGCCGCT GATTATCGAC GCGGGCTTGG GTTTGCCTTC
601
    CCAAGCGGCA CAAGTGATGG AATGGGGTTT TGACGGCGTA TTGTTAAACA
651
701 CCGCCGTTTC CCGCAGCGGC GACCCCGTCA ACATGGCGCG CGCCTTCGCA
751 CTCGCCGTCG AATCCGGACG GCTGGCATTT GAAGCCGGGC CGGTCGAAGC
    GCGAACCAAA GCCCAAGCCA GCACGCCGAC AGTCGGACAA CCGTTTTGGC
851 ATTCGGCGGA ATATTGA
```

### This corresponds to the amino acid sequence <SEQ ID 464; ORF 126.ng>:

g126.pep

- 1 MPSETPKARR RLSDGIASDN HTKESIMLTL YGETFPSRLL LGTAAYPTPE 51 ILKQSVRTAR PAMITVSLRR TGCGGEAHGQ GFWSLLQETG VPVLPNTAGC
- 101 QSVQEAVTTA QMAREVFETD WIKLELIGDD DTLQPDVFQL VEAAEILIKD
- 151 GFKVLPYCTE DLIACRRLLD AGCQALMPWA APIGTGLGAV HAYALKILRE
- 201 RLPDTPLIID AGLGLPSQAA QVMEWGFDGV LLNTAVSRSG DPVNMARAFA
- 251 LAVESGRLAF EAGPVEARTK AQASTPTVGQ PFWHSAEY*

## The following partial DNA sequence was identified in N. meningitidis <SEQ ID 465>:

```
..CACTATACAA AGGAACCCAT TATGCTCACC CTATACGGCG AAACTTTCCC
  1
       CTCGCGGCTG CTGCTCGGCA CGGCTGCCTA CCCGACCCCC GAAATCCTCA
 51
       AACAATCCAT CCAAACCGCC CAGCCTGCGA TGATTACCGT CTCGCTGCGC
101
       CGCGCGGGAA GCGGCGCGA GGCGCACGGT CAGGGGTTTT GGTCGCTGCT
       TCAAGAAACC GGCGTTCCCG TCCTGCCGAA CACGGCAGGC TGCCAAAGCG
       TGCAGGAAGC GGTAACGACG GCGCAAATGG CGCGCGAAGT GTTTGAAACC
       GATTGGATAA AATTGGAACT CATCGGAGAT GACGACACCT TGCAGCCGGA
301
       TGTGTTCCAG CTTGTCGAAG CGGCGGAAAT CCTGATTAAA GACGGCTTCA
351
       AAGTGCTGCC TTATTGCACC GAAGACCTGA TTGCCTGCCG CCGCCTGCTC
401
451
       GACGCGGGCT GTCAGGCGTT GATGCCGTGG GCGGCCCCGA TCGGCACGGG
       TTTGGGCGCG GTTCACGCCT ACGCGTTGAA CGTCCTGCGC GAACGCCTGC
       CCGACACGCC GCTGATTATC GACGCGGGCT TGGGTTTGCC CTCACAGGCG
551
       GCACAAGTGA TGGAATGGGG CTTTGACGGC GTGCTTTTGA ATACTGCCGT
601
       TTCCCGCAGC GGCGATCCGG TCAATATGGC ACGCGCCTTC GCACTCGCCG
651
701
       TCGAATCCGG ACGGCTGGCA TTTGAAGCCG GACCGGTCGA AGCACGCGAC
      AAAGCGCAAG CCAGCACGCC GACAGTCGGA CAACCGTTTT GGCATTCGGC
751
801
       GGAATATTGA
```

## This corresponds to the amino acid sequence <SEQ ID 466; ORF 126>: m126.pep (partial)

1 ..HYTKEPIMLT LYGETFPSRL LLGTAAYPTP EILKQSIQTA QPAMITVSLR 51 RAGSGGEAHG QGFWSLLQET GVPVLPNTAG CQSVQEAVTT AQMAREVFET

101	DWIKLELIGD	DDTLQPDVFQ	LVEAAEILIK	DGFKVLPYCT	EDLIACRRLL
151	DAGCQALMPW	AAPIGTGLGA	VHAYALNVLR	ERLPDTPLII	DAGLGLPSQA
201	AQVMEWGFDG	VLLNTAVSRS	GDPVNMARAF	ALAVESGRLA	FEAGPVEARD
251	KAQASTPTVG	QPFWHSAEY*			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae

ORF 126 shows 95.9% identity over a 269 aa overlap with a predicted ORF (ORF 126.ng) from N. gonorrhoeae:

m126/g126

m126.pep		нутк	10 EPIMLTLYGE	20 TFPSRLLLGT	30 AAYPTPEILK	40 OSTOTAO
		::11				::  :
g126	MPSETPKARRRLSD	GIASDNHTK	ESIMLTLYGE'	TFPSRLLLGT	AAYPTPEILK	QSVRTAR
	10	20	30	40	50	60
	50	60	70	80	90	100
m126.pep	PAMITVSLRRAGSG	GEAHGQGFW	SLLQETGVPV	LPNTAGCOSV	OEAVTTAOMA	REVFETD
				11111111	ĪHHHĪH	111111
g126	PAMITVSLRRTGCG	GEAHGQGFW	SLLQETGVPV	LPNTAGCOSV	DEAVTTAOMA	REVEETD
	70	80	90	100	110	120
	110	120	130	140	150	160
m126.pep	WIKLELIGDDDTLQ					
zo.pcp	1111111111111				HURRILIDAGU	QALMPWA
g126	WIKLELIGDDDTLQ		APTITITITE PR			ONIMPER
9120	130	140	150	160	170	QALMPWA. 180
	130	140	150	100	170	180
	170	180	190	200	210	220
m126.pep	APIGTGLGAVHAYA	LNVLRERLP	DTPLIIDAGLO	GLPSQAAQVM	EWGFDGVLLN	TAVSRSG
		1::				
g126	APIGTGLGAVHAYA			GLPSQAAQVM	EWGFDGVLLN	TAVSRSG
	190	200	210	220	230	240
	230	240	250	260	270	
m126.pep	DPVNMARAFALAVE	SGRLAFEAG	PVEARDKAQAS	STPTVGOPFWI	HSAEYX	
		111111111				
g126	DPVNMARAFALAVE	SGRLAFEAG	PVEARTKAOAS	STPTVGOPFWI	HSAEYX	
-	250	260	270	280	<b></b>	

### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 467>: a126

126.seq					
1	TTGTTAATCC	ACTATACAAA	GGAACCCATT	ATGCTCACCC	TGTACAGCGA
51	AACTTTCCCT	TCGCGGCTGC	TGCTCGGCAC	AGCCGCCTAC	CCGACCCCTG
101	AAATCCTCAA	ACAATCCGTC	CGAACCGCCC	GGCCCGCGAT	GATTACCGTC
151	TCGCTGCGCC	GCGCGGGATG	CGGCGGCGAG	GCGCACGGTC	AGGGGTTTTG
201	GTCGCTGCTT	CAAGAAACCG	GCGTTCCCGT	CCTGCCGAAC	ACGGCAGGCT
251	GCCAAAGCGT	GCAGGAAGCG	GTAACGACGG	CGCAAATGGC	GCGCGAAGTG
301	TTTGAAACCG	ATTGGATTAA	ACTCGAACTC	ATCGGCGACG	ACGACACCTT
351	GCAGCCGGAT	GTGTTCCAAC	TTGTCGAAGC	GGCGGAAATC	CTGATTAAAG
401	ACGGCTTCAA	AGTGCTGCCT	TATTGCACCG	AAGACCTGAT	TGCCTGCCGC
451	CGCCTGCTCG	ACGCGGGCTG	TCAGGCGTTG	ATGCCGTGGG	CGGCCCCGAT
501	CGGCACGGGT	TTGGGCGCGG	TTCACGCCTA	CGCGTTGAAC	GTCCTGCGCG
551	AACGCCTGCC	CGACACGCCG	CTGATTATCG	ACGCGGGCTT	GGGTTTGCCC
601	TCACAGGCGG	CACAAGTGAT	GGAATGGGGC	TTTGACGGCG	TGCTTTTGAA
651		TCCCGCAGCG		CAATATGGCA	
701	CACTCGCCGT	CGAATCCGGA	CGGCTGGCAT	TTGAAGCCGG	ACCGGTCGAA
751	GCACGCGACA	AAGCGCAAGC	CAGCACGCCG	ACAGTCGGAC	AACCGTTTTG
801	GCATTCGGCG	GAATATTGA			

a126

This corresponds to the amino acid sequence <SEQ ID 468; ORF 126.a>: a126.pep LLIHYTKEPI MLTLYSETFP SRLLLGTAAY PTPEILKQSV RTARPAMITV 51 SLRRAGCGGE AHGQGFWSLL QETGVPVLPN TAGCQSVQEA VTTAQMAREV FETDWIKLEL IGDDDTLQPD VFQLVEAAEI LIKDGFKVLP YCTEDLIACR 101 RLLDAGCQAL MPWAAPIGTG LGAVHAYALN VLRERLPDTP LIIDAGLGLP 151 SQAAQVMEWG FDGVLLNTAV SRSGDPVNMA RAFALAVESG RLAFEAGPVE 201 ARDKAQASTP TVGQPFWHSA EY* 251 m126/a126 98.1% identity in 269 aa overlap 20 10 30 40 50 HYTKEPIMLTLYGETFPSRLLLGTAAYPTPEILKQSIQTAQPAMITVSLRRAGSGGE m126.pep a126  $\verb|LLIHYTKEPIMLTLYSETFPSRLLLGTAAYPTPEILKQSVRTARPAMITVSLRRAGCGGE|$ 30 40 60 80 90 100 AHGQGFWSLLQETGVPVLPNTAGCQSVQEAVTTAQMAREVFETDWIKLELIGDDDTLQPD m126.pep AHGQGFWSLLQETGVPVLPNTAGCQSVQEAVTTAQMAREVFETDWIKLELIGDDDTLQPD a126 70 80 90 100 120 130 140 150 160 VFQLVEAAEILIKDGFKVLPYCTEDLIACRRLLDAGCQALMPWAAPIGTGLGAVHAYALN m126.pep a126 VFQLVEAAEILIKDGFKVLPYCTEDLIACRRLLDAGCQALMPWAAPIGTGLGAVHAYALN 130 140 150 160 180 190 200 210 220 230 VLRERLPDTPLIIDAGLGLPSQAAQVMEWGFDGVLLNTAVSRSGDPVNMARAFALAVESG m126.pep a126 VLRERLPDTPLIIDAGLGLPSQAAQVMEWGFDGVLLNTAVSRSGDPVNMARAFALAVESG 190 200 210 220 230 240 240 250 260 m126.pep RLAFEAGPVEARDKAQASTPTVGQPFWHSAEYX 

# The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 469>: g126-1.seq

RLAFEAGPVEARDKAQASTPTVGQPFWHSAEYX

260

270

250

	•				
1	ATGCTCACCC	TGTACGGCGA	AACTTTCCCT	TCGCGGCTGC	TGCTCGGCAC
51	GGCCGCCTAC	CCGACCCCTG	AAATCCTCAA	ACAATCCGTC	CGAACCGCCC
101	GGCCCGCGAT	GATTACCGTC	TCGCTGCGCC	GCACGGGATG	CGGCGGCGAG
151	GCGCACGGTC	AGGGGTTTTG	GTCGCTGCTT	CAAGAAACCG	GCGTTCCCGT
201	CCTGCCGAAC	ACGGCAGGCT	GCCAAAGCGT	GCAGGAAGCG	GTAACGACGG
251	CGCAAATGGC	GCGCGAAGTG	TTTGAAACCG	ATTGGATAAA	ATTGGAACTC
301	ATCGGCGACG	ACGACACCTT	GCAGCCGGAC	GTGTTCCAAC	TCGTCGAAGC
351	GGCGGAAATC	CTGATTAAAG	ACGGCTTCAA	AGTGCTGCCT	TATTGCACCG
401	AAGACCTGAT	TGCCTGCCGC	CGCCTGCTCG	ATGCGGGCTG	TCAGGCGTTG
451	ATGCCGTGGG	CGGCTCCCAT	CGGCACGGGT	TTGGGGGCGG	TTCACGCCTA
501	TGCGCTCAAA	ATCCTGCGCG	AACGCCTGCC	CGACACGCCG	CTGATTATCG
551	ACGCGGGCTT	GGGTTTGCCT	TCCCAAGCGG	CACAAGTGAT	GGAATGGGGT
601	TTTGACGGCG	TATTGTTAAA	CACCGCCGTT	TCCCGCAGCG	GCGACCCCGT
651	CAACATGGCG	CGCGCCTTCG	CACTCGCCGT	CGAATCCGGA	CGGCTGGCAT
701	TTGAAGCCGG	GCCGGTCGAA	GCGCGAACCA	AAGCCCAAGC	CAGCACGCCG
751	ACAGTCGGAC	AACCGTTTTG	GCATTCGGCG	GAATATTGA	•
	51 101 151 201 251 301 351 401 451 501 551 601 651 701	51 GGCCGCCTAC 101 GGCCCGCGAT 151 GCGCACGGTC 201 CCTGCCGAAC 251 CGCAAATGGC 301 ATCGGCGACG 351 GGCGGAAATC 401 AAGACCTGAT 451 ATGCCGTGGG 501 TGCGCTCAAA 551 ACGCGGCTT 601 TTTGACGGCG 701 TTGAAGCCGG	51 GGCCGCCTAC CCGACCCCTG 101 GGCCCGCGAT GATTACCGTC 151 GCGCACGGTC AGGGGTTTTG 201 CCTGCCGAAC ACGGCAGGTC 301 ATCGGCGACG ACGACACCTT 351 GGCGGAAATC CTGATTAAAG 401 AAGACCTGAT TGCCTGCCGC 451 ATGCCGTGGG CGGCTCCCAT 501 TGCGCTCAAA ATCCTGCCCG 551 ACGCGGGCTT GGGTTTGCCT 601 TTTGACGGCG TATTGTTAAA 651 CAACATGGCG CGCGCCTTCG 701 TTGAAGCCGG GCCGGTCGAA	51 GGCCGCCTAC CCGACCCCTG AAATCCTCAA 101 GGCCCGCGAT GATTACCGTC TCGCTGCGCC 151 GCGCACGGTC AGGGGTTTTG GTCGCTGCTT 201 CCTGCCGAAC ACGGCAGGTT GCCAAAGCGT 251 CGCAAATGGC GCGCGAAGTG TTTGAAACCG 301 ATCGGCGACG ACGACCACCTT GCAGCCGGAC 351 GGCGGAAATC CTGATTAAAG ACGGCTTCAA 401 AAGACCTGAT TGCCTGCCGC CGCCTGCTCG 451 ATGCCGTGGG CGGCTCCCAT CGGCACGGGT 501 TGCGCTCAA ATCCTGCCGC AACGCCTGCC 551 ACGCGGGCTT GGGTTTGCCT TCCCAAGCGG 601 TTTGAAGCCGG CGCGCCTTCG CACTCGCCGT 701 TTGAAGCCGG GCCGGTCGAA GCGCGAACCA	51 GGCCGCTAC CCGACCCCTG AAATCCTCAA ACAATCCGTC 101 GGCCCGCGAT GATTACCGTC TCGCTGCGCC GCACGGGATG 151 GCGCACGGTC AGGGGTTTTG GTCGCTGCTT CAAGAAACCG 201 CCTGCCGAAC ACGGCAAGTG GCCAAAGCGT GCAGGAAGCG 251 CGCAAATGGC GCGGAAGTG TTTGAAACCG ATTGGATAAC 301 ATCGGCGACG ACGACACCTT GCAGCCGGAC GTGTTCCAAC 351 GGCGGAAATC CTGATTAAAG ACGGCTTCAA AGTGCTGCCT 401 AAGACCTGAT TGCCTGCCGC CGCCTGCTCG ATGCGGGCTG 451 ATGCCGTGGG CGGCTCCCAT CGGCACGGGT TTGGGGGCGG 501 TGCGCTCAAA ATCCTGCGCG AACGCCTGCC CGACACGCCG 551 ACGCCGGCTT GGGTTTGCCT TCCCAAGCGG CACAAGTGAT 601 TTTGACGGCG TATTGTTAAA CACCGCCGTT TCCCGCAGCG 651 CAACATGGCG CGCGCCTTCG CACTCGCCGT CGAATCCGGA 701 TTGAAGCCGG GCCGGTCGAA GCGCGAACCA AAGCCCAAGC

This corresponds to the amino acid sequence <SEQ ID 470; ORF 126-1.ng>: g126-1.pep

¹ MLTLYGETFP SRLLLGTAAY PTPEILKQSV RTARPAMITV SLRRTGCGGE

```
51 AHGQGFWSLL QETGVPVLPN TAGCQSVQEA VTTAQMAREV FETDWIKLEL
101 IGDDDTLQPD VFQLVEAAEI LIKDGFKVLP YCTEDLIACR RLLDAGCQAL
151 MPWAAPIGTG LGAVHAYALK ILRERLPDTP LIIDAGLGLP SQAAQVMEWG
201 FDGVLLNTAV SRSGDPVNMA RAFALAVESG RLAFEAGPVE ARTKAQASTP
251 TVGQPFWHSA EY*
```

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 471>: m126-1.seq

```
1 ATGCTCACCC TATACGGCGA AACTTTCCCC TCGCGGCTGC TGCTCGGCAC
    GGCTGCCTAC CCGACCCCG AAATCCTCAA ACAATCCATC CAAACCGCCC
    AGCCTGCGAT GATTACCGTC TCGCTGCGCC GCGCGGGAAG CGGCGGCGAG
101
    GCGCACGGTC AGGGGTTTTG GTCGCTGCTT CAAGAAACCG GCGTTCCCGT
151
201 CCTGCCGAAC ACGGCAGGCT GCCAAAGCGT GCAGGAAGCG GTAACGACGG
    CGCAAATGGC GCGCGAAGTG TTTGAAACCG ATTGGATAAA ATTGGAACTC
251
301 ATCGGAGATG ACGACACCTT GCAGCCGGAT GTGTTCCAGC TTGTCGAAGC
    GGCGGAAATC CTGATTAAAG ACGGCTTCAA AGTGCTGCCT TATTGCACCG
351
401 AAGACCTGAT TGCCTGCCGC CGCCTGCTCG ACGCGGGCTG TCAGGCGTTG
    ATGCCGTGGG CGGCCCCGAT CGGCACGGGT TTGGGCGCGG TTCACGCCTA
451
    CGCGTTGAAC GTCCTGCGCG AACGCCTGCC CGACACGCCG CTGATTATCG
    ACGCGGGCTT GGGTTTGCCC TCACAGGCGG CACAAGTGAT GGAATGGGGC
551
    TTTGACGGCG TGCTTTTGAA TACTGCCGTT TCCCGCAGCG GCGATCCGGT
601
    CAATATGGCA CGCGCCTTCG CACTCGCCGT CGAATCCGGA CGGCTGGCAT
651
701 TTGAAGCCGG ACCGGTCGAA GCACGCGACA AAGCGCAAGC CAGCACGCCG
   ACAGTCGGAC AACCGTTTTG GCATTCGGCG GAATATTGA
```

This corresponds to the amino acid sequence <SEQ ID 5; ORF 126-1>: m126-1.pep

```
1 MLTLYGETFP SRLLLGTAAY PTPEILKQSI QTAQPAMITV SLRRAGSGGE
51 AHGQFWSLL QETGVPVLPN TAGCQSVQEA VTTAQMAREV FETDWIKLEL
101 IGDDDTLQPD VFQLVEAAEI LIKDGFKVLP YCTEDLIACR RLLDAGCQAL
151 MPWAAPIGTG LGAVHAYALN VLRERLPDTP LIIDAGLGLP SQAAQVMEWG
201 FDGVLLNTAV SRSGDPVNMA RAFALAVESG RLAFEAGPVE ARDKAQASTP
251 TVGQPFWHSA EY*
```

m126-1/g126-1 96.9% identity in 262 aa overlap 20 40 10 30 MLTLYGETFPSRLLLGTAAYPTPEILKQSIQTAQPAMITVSLRRAGSGGEAHGQGFWSLL m126-1.pep MLTLYGETFPSRLLLGTAAYPTPEILKQSVRTARPAMITVSLRRTGCGGEAHGQGFWSLL g126-1 20 30 40 10 80 QETGVPVLPNTAGCQSVQEAVTTAQMAREVFETDWIKLELIGDDDTLQPDVFQLVEAAEI m126-1.pep g126-1 QETGVPVLPNTAGCQSVQEAVTTAQMAREVFETDWIKLELIGDDDTLQPDVFQLVEAAEI 70 80 90 100 150 160 170  $\verb|Likdgfkvlpyctedliacrrlldagcqalmpwaapigtglgavhayalnvlrerlpdtp|$ m126-1.pep g126-1 LIKDGFKVLPYCTEDLIACRRLLDAGCQALMPWAAPIGTGLGAVHAYALKILRERLPDTP 130 140 150 160 180 190 200 210 220 230

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 473>: a126-1.seq

- 1 ATGCTCACCC TGTACAGCGA AACTTTCCCT TCGCGGCTGC TGCTCGGCAC
- 51 AGCCGCCTAC CCGACCCCTG AAATCCTCAA ACAATCCGTC CGAACCGCCC
- 101 GGCCCGCGAT GATTACCGTC TCGCTGCGCC GCGCGGGATG CGGCGGCGAG

BNSDOCID: <WO___9957280A2_j_>

Sept. 18



```
GCGCACGGTC AGGGGTTTTG GTCGCTGCTT CAAGAAACCG GCGTTCCCGT
201
     CCTGCCGAAC ACGGCAGGCT GCCAAAGCGT GCAGGAAGCG GTAACGACGG
     CGCAAATGGC GCGCGAAGTG TTTGAAACCG ATTGGATTAA ACTCGAACTC
251
     ATCGGCGACG ACGACACCTT GCAGCCGGAT GTGTTCCAAC TTGTCGAAGC
301
351
     GGCGGAAATC CTGATTAAAG ACGGCTTCAA AGTGCTGCCT TATTGCACCG
     AAGACCTGAT TGCCTGCCGC CGCCTGCTCG ACGCGGGCTG TCAGGCGTTG
401
     ATGCCGTGGG CGGCCCCGAT CGGCACGGGT TTGGGCGCGG TTCACGCCTA
CGCGTTGAAC GTCCTGCGCG AACGCCTGCC CGACACGCCG CTGATTATCG
451
501
     ACGCGGGCTT GGGTTTGCCC TCACAGGCGG CACAAGTGAT GGAATGGGGC
551
     TTTGACGGCG TGCTTTTGAA TACTGCCGTT TCCCGCAGCG GCGATCCGGT
601
     CAATATGGCA CGCGCCTTCG CACTCGCCGT CGAATCCGGA CGGCTGGCAT
651
     TTGAAGCCGG ACCGGTCGAA GCACGCGACA AAGCGCAAGC CAGCACGCCG
701
751 ACAGTCGGAC AACCGTTTTG GCATTCGGCG GAATATTGA
```

#### This corresponds to the amino acid sequence <SEQ ID 474; ORF 126-1.a>: a126-1.pep

- MLTLYSETFP SRLLLGTAAY PTPEILKQSV RTARPAMITV SLRRAGCGGE 1 51
- AHGQGFWSLL QETGVPVLPN TAGCQSVQEA VTTAQMAREV FETDWIKLEL 101
- IGDDDTLQPD VFQLVEAAEI LIKDGFKVLP YCTEDLIACR RLLDAGCQAL MPWAAPIGTG LGAVHAYALN VLRERLPDTP LIIDAGLGLP SQAAQVMEWG 151
- 201 FDGVLLNTAV SRSGDPVNMA RAFALAVESG RLAFEAGPVE ARDKAQASTP
- TVGOPFWHSA EY*

WO 99/57280

#### a126-1/m126-1 98.1% identity in 262 aa overlap

```
10
                          20
                                   30
           MLTLYSETFPSRLLLGTAAYPTPEILKQSVRTARPAMITVSLRRAGCGGEAHGQGFWSLL
 a126-1.pep
           инизининининий::пзинини ининин
m126-1
           MLTLYGETFPSRLLLGTAAYPTPEILKQSIQTAQPAMITVSLRRAGSGGEAHGQGFWSLL
                                  30
                                           40
                                                   50
                  70
                          80
                                  90
                                          100
                                                  110
           QETGVPVLPNTAGCQSVQEAVTTAQMAREVFETDWIKLELIGDDDTLQPDVFQLVEAAEI
a126-1.pep
           m126-1
           QETGVPVLPNTAGCQSVQEAVTTAQMAREVFETDWIKLELIGDDDTLQPDVFQLVEAAEI
                  70
                          80
                                  90
                                                  110
                 130
                         140
                                 150
                                          160
           {	t LIKDGFKVLPYCTEDLIACRRLLDAGCQALMPWAAPIGTGLGAVHAYALNVLRERLPDTP}
a126-1.pep
           m126-1
           LIKDGFKVLPYCTEDLIACRRLLDAGCQALMPWAAPIGTGLGAVHAYALNVLRERLPDTP
                 130
                         140
                                 150
                                         160
                                                  170
                                                          180
                 190
                         200
                                 210
                                         220
                                                  230
a126-1.pep
           LIIDAGLGLPSQAAQVMEWGFDGVLLNTAVSRSGDPVNMARAFALAVESGRLAFEAGPVE
           m126-1
           LIIDAGLGLPSQAAQVMEWGFDGVLLNTAVSRSGDPVNMARAFALAVESGRLAFEAGPVE
                 190
                         200
                                 210
                                         220
                 250
                         260
a126-1.pep
           ARDKAQASTPTVGQPFWHSAEYX
           11111111111111111111111111
m126-1
           ARDKAQASTPTVGQPFWHSAEYX
                250
                        260
```

#### The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 475>: g127.seq

ATGGAAATAT	GGAATATGTT	GAACACTTGG	CCCGATGCCG	TCCCCATACC
CGCGGAGGCG	GCCGAATCCG	TGGCGGCGGT	CGCGGCTTTC	COCCGATACG
GCGCCCTTCT	GTTGAATATC	CACTTCAGAC	CCCATCCCCA	CIGCIGGCGC
GAAAGCAAGC	CCCCCTTTTT	CCTTCCCACC	GGCATCCGGA	TTTCGGCATC
GGTGCTGTTT	TCCCTCCCAT	TTT TOTCAGE	CGCAATATAA	CGCTGCTTTT
CTTTCTCCAT	CTTTCCCCCCC	TIATCTGGTC	GGCGCAAATT	CAAACGCTGG
CTCATTATCT	GITIGCGGTG	GCGGCGGCGG	TCGTCGTGGC	GACAAAAGAA
CTGATTAIGI	GTCTGTCGGG	CAGTATTTTA	aggtctGCCA	CCCAGCAATA
CICGGICGGC	GACTATATCG	AAATCAACGG	CCTGCGCGGG	CCCCTCCTCC
ACATCAATCT	GTTGAACACG	CTGATGATGC	AGGTCGGTCC	CAACCCCCCCC
GTCGGACAGC	TTGCGGGAAC	CACCGTTTCT	TTCCCCAACA	CCCMCMMcmm
GAGCCACCCC	GTGCGCCGCG	ACAATATTTT	GGGCGACTAT	GTCATCCATA
	CGCGGAGGCG GCGCCCTTCT GAAAGCAAGC GGTGCTGTTT CTTTGTCGAT CTGATTATGT CTCGGTCGGC ACATCAATCT GTCGGACAGC	GCGCGGAGGCG GCCGAATCCG GCGCCCTTCT GTTGAATATC GAAAGCAAGC GGCGGTTTTT GGTGCTGTTT TCGCTGGCAT CTTTGTCGAT GTTTGCGGTG CTGATTATGT GTCTGTCGGG CTCGGTCGGC GACTATATCG ACATCAATCT GTTGAACACG GTCGGACAGC TTGCGGGAAC	CGCGGAGGCG GCCGAATCCG TGGCGGCGGT GCGCCCTTCT GTTGAATATC CACTTCAGAC GAAAGCAAGC GGCGGTTTTT GGTTGCCAGC GGTGCTGTTT TCGCTGGCAT TTATCTGGTC CTTTGTCGAT GTTTGCGGTG GCGGCGGCGG CTGATTATGT GTCTGTCGGG CAGTATTTTA CTCGGTCGGC GACTATATCG AAATCAACGG ACATCAATCT GTTGAACACG CTGATGATGC GTCGGACAGC TTGCGGGAAC CACCGTTTCT	ATGGAAATAT GGAATATGTT GAACACTTGG CCCGATGCCG CGCGGAGGCG GCCGAATCCG TGGCGGCGGT CGCGGCTTTG GCGCCCTTCT GTTGAATATC CACTTCAGAC GGCAATCCGGA GAAAGCAAGC GGCGGTTTTT GGTTGCCAGC CGCAATATAA GGTGCTGTTT TCGCTGGCAT TTATCTGGTC GGCGCAAATT CTTTGTCGAT GTTTGCGGGT CAGTATTTTA aggtctGCCA CTCGGTCGGC GACTATATCG AAATCAACGG CCTGCGCGGG ACATCAATCT GTTGAACACG CTGATGATGC AGGTCGGTCC GTCGGACAGC TTGCGGGAAC CACCGTTTCT TTCCCCAACA GAGCCACCCC GTGCGCCGC ACAATATTTT GGGCGACTAT

```
551 CGGTCGAAAT CCCCGTTCCC ATCCATTTGG ATTCGGATGA AGCCGTATGC
              CGTCTGAAAG CCGTACTCGA GCCCTTGTGC GCGCCCTACA TCCCCGCCAT
              TCAGCGGTAT TTGGAAAACG TGCAGGCGGA AAAACTGTTT ATCACGCCCG
               CCGCCAGGCC GCGCGTTACC CGCGTACCGT ACGACGACAA GGCATACCGC
          751 ATCATCGTCC GCTTCGCCTC CCCCGTTTCA AAGCGGCTGG AAATCCAACA
              GGCGGTTATG GACGAATTTT TGCGCGTACA ATACCGCCTG TTAAATCATC
          801
          851
               CCGCCGgctc cgAAACACTT TAA
This corresponds to the amino acid sequence <SEQ ID 476; ORF 127.ng>:
     g127.pep
               MEIWNMLNTW PDAVPIRAEA AESVAAVAAL LLARALLLNI HFRRHPDFGI
            1
           51
              ESKRRFLVAS RNITLLLVLF SLAFIWSAQI QTLALSMFAV AAAVVVATKE
              LIMCLSGSIL RSATOOYSVG DYIEINGLRG RVVDINLLNT LMMOVGPNPL
          151 VGQLAGTTVS FPNSLLLSHP VRRDNILGDY VIHTVEIPVP IHLDSDEAVC
          201 RLKAVLEPLC APYIPAIQRY LENVQAEKLF ITPAARPRVT RVPYDDKAYR
          251 IIVRFASPVS KRLEIQQAVM DEFLRVQYRL LNHPAGSETL *
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 477>:
     m127.seq
              ATGGAAATAT GGAATATGTT GGACACTTGG CTCGGTGCCG TCCCGATACG
            1
           51
              TGCGGAGGCG GTCGAATCCG TGGCGGCGGT TGCGGCTTTG CTGCTGGCGC
              GCGCCCTTCT GTTGAATATC CACTTCAAAC GGCATCCGGA TTTCGGCATC
          151 GAAAGCAAGC GGCGGTTTTT GGTTGCCAGC CGCAATATAA CGCTGCTTTT
          201 GGTGCTGTTT TCGCTGGCAT TTATCTGGTC GGCGCAAATC CAAACGCTGG
          251 CTTTGTCGAT GTTTGCGGTG GCGGCGGCGG TCGTCGTGGC GACGAAGGAA
          301 CTGATTATGT GTCTGTCGGG CAGTATTTTA AGGTCTGCCA CCCAGCAATA
          351 CTCGGTCGGC GACTATATCG AAATCAACGG CCTGCGCGGG CGCGTGGTCG
          401 ACATCAACCT GTTGAACACG CTGATGATGC AGGTCGGTCC GAACCCCTTG
          451 GTCGGACAGC TTGCGGGAAC CACCGTTTCT TTCCCCAACA GCCTGTTGTT
          501 GAGCCACCCC GTGCGCCGCG ACAATATTTT GGGCGACTAT GTCATCCATA
          551 CGGTCGAAAT CCCCGTTCCC ATCCATTTGG ATTCGGATGA AGCCGTATGC
          601 CGTCTGAAAG CCGTACTCGA GCCCTTGTGC GCGCCCTACA TCCCCGCCAT
          651 CCAACGGSAT TTGGAAAACG TGCAGGCGGA AAAACTGTTT ATCACGCCCG
              CCGCCAGACC GCGCGTTACC CGCGTGCCGT ACGATGACAA GGCATACCGC
          751 ATCATCGTCC GCTTCGCTTC CCCCGTTTCA AAGCGGCTGG AAATCCAACA
              GGCGGTTATG GACGAATTTT TGCGCGTACA ATACCGCCTG TTAAATCACC
              CCGCCGGCTC CGAAACACTT TAA
          851
This corresponds to the amino acid sequence <SEQ ID 478; ORF 127>:
     m127.pep
              MEIWNMLDTW LGAVPIRAEA VESVAAVAAL LLARALLLNI HFKRHPDFGI
           51
              ESKRRFLVAS RNITLLLVLF SLAFIWSAQI QTLALSMFAV AAAVVVATKE
              LIMCLSGSIL RSATQQYSVG DYIEINGLRG RVVDINLLNT LMMQVGPNPL
              VGQLAGTTVS FPNSLLLSHP VRRDNILGDY VIHTVEIPVP IHLDSDEAVC
              RLKAVLEPLC APYIPAIQRX LENVQAEKLF ITPAARPRVT RVPYDDKAYR
              IIVRFASPVS KRLEIQQAVM DEFLRVQYRL LNHPAGSETL *
Computer analysis of this amino acid sequence gave the following results:
Homology with a predicted ORF from N. gonorrhoeae
ORF 127 shows 97.9% identity over a 290 aa overlap with a predicted ORF (ORF 127.ng)
from N. gonorrhoeae:
     m127/g127
                         10
                                   20
                                            30
                                                      40
                                                                          60
    m127.pep
                 MEIWNMLDTWLGAVPIRAEAVESVAAVAALLLARALLLNIHFKRHPDFGIESKRRFLVAS
                 MEIWNMLNTWPDAVPIRAEAAESVAAVAALLLARALLLNIHFRRHPDFGIESKRRFLVAS
     g127
                         10
                                   20
                                            30
                                                      40
                                   80
                                            . 90
                                                     100
                                                               110
                                                                         120
                 RNITLLLVLFSLAFIWSAQIQTLALSMFAVAAAVVVATKELIMCLSGSILRSATQQYSVG
    m127.pep
                 g127
                 RNITLLLVLFSLAFIWSAQIQTLALSMFAVAAAVVVATKELIMCLSGSILRSATQQYSVG
```

BNSDOCID: <WO___9957280A2_l_>

	70	80	90	100	110	120
	130	140	150			120
m127.pep	DYIEINGLRGRVVD	INLLNTLM	OVGPNPLAGOI	160 AGTTVSFPNS	170 SLLLSHPVRRF	180 NTLCDY
g127	1   1   1   1   1   1   1   1   1   1	1111111		111111111		
<b>3</b>	DYIEINGLRGRVVD 130	140	IQVGPNPLVGQL 150	AGTTVSFPNS 160	LLLSHPVRRD 170	DNILGDY
	100	222			170	180
m127.pep		200 SDEAVCRLK	210 AVLEPLCAPYT	220	230	240
q127			1111111	1111 111	111111111	111111
9127	VIHTVEIPVPIHLD	SDEAVCRLK 200	AVLEPLCAPYI 210	PAIQRYLENV	QAEKLFITPA	ARPRVT
			210	220	230	240
m127.pep	250 RVPYDDKAYRIIVRI	260 FASDVSKDT:	270	280	290	
			1 [ 1 [ 1 [ 1 [ 1 [ 1 ] ] ]	11111111	111111	
g127	RVPYDDKAYRIIVRI 250	ASPVSKRL	EIQQAVMDEFL	RVQYRLLNHP	AGSETLX	
		260	270	280	290	
The following p	partial DNA sequence	was ident	tified in N. m	eningitidis	<seo 4<="" id="" td=""><td>l79&gt;·</td></seo>	l79>·
a127.seq 1						,,,,,
51	ATGGAAATAT GGAATAT TGCGGAGGCG GTCGAAT	CCG TGGC	GTGGT CGCGG	COMPAC CHCC	TO COOO	
101 151	GCGCCCTTCT GTTGAAT	ATC CACTI	יכאאאר ההראיז	CCCCA mmmc	2000000	
201	GHANGCHAGC GGCGGTT	TIT GGTTC	CCAGC CGCAA	ጥለጥለአ ሶርርብ	COMMON	
251		GTG GCGGC	CIGGIC GGCGC	CAATC CAAF	CGCTGG	
301	CIGALIATET GTCTGTC	GGG CAGCA	עדידידים בכידיר	TICCON COOM	CONNEN	
351	CICGGICGGC GACTATA	TCG AAATC	'AACGG CCTGC	CCCCC CCCC	mccmee.	
401 451	ACAICAACCI GITGAAC	ACG CTGAT	'GATGC გვვლი	CCTCC CNNC	CCCmmc	
501	GICGGACAGC TTGCGGG	AAC CACCG	ጥጥጥርጥ ጥጥርርር	ראארא כככה	CMMCMM	
551		CCC ATCCA	ATTTT GGGCG	ACTAC GTCA	TCCATA	
601	CGTCTGAAAG CCGTACTO	CGA GCCCT	TGTGC GCGCC	GATGA AGCC	GTATGC	
651	CCAACGGCAT TIGGAAAA	ACG TGCAG	GCGGA AAAAC	עכיישע אשכיא	000000	
701	CCGCCAAACC GCGCGTT	ACC CGCGT	GCCGT ACGAT	CACAA CCCA	m » o o o o	
751 801	ATCATOGTOC GCTTCGC	JI'O CCCCG	TTTCA AAGCG	מכיייכר אאאיי	CC2 2 C2	
851	GGCGGTTATG GACGAATT	TTT TGCGC	GTACA ATACC	GCCTG TTAA	ATTACC	
This sames 1						
al27.pep	s to the amino acid see	quence <s< td=""><td>SEQ ID 480;</td><td>ORF 127.a</td><td>&gt;:</td><td></td></s<>	SEQ ID 480;	ORF 127.a	>:	
a127.pep	METUNMI DEM I CAMPIDA					
51	MEIWNMLDTW LGAVPIRA ESKRRFLVAS RNITLLLV	TE SLAFT	VVAAL LLARAI	LLINI HFKRI	IPDFGI	
101	TINCTOGOIT KOMIOOAS	IVG DYTETN	JGT.RG DUMINTA	IT TAIM TAGACOT	100000	
151	AGGTWGIIA2 LENSTTTS	HP VRRDN	こうしん ひょうかいち	TINT CITTE		
201 251	VIVAATERIC ASITSATO	KH LENVO	AFKIF TTDAAL	זערונות שווססי	DKAYR	
251	IIVRFASPVS KRLEIQQA	VM DEFLR	OYRL LNYPAG	SSETL *		
m127/a127 98.	6% identity in 290 aa	overlap				
m127.pep	10	20	30	40	50	60
miz/.pep	MEIWNMLDTWLGAVPI	RAEAVESVA	AVAALLLARAL	LLNIHFKRHF	DECTRONS	
a127		RAEAVESVA		INTERPER	111111111	1111
	10	20	30	40	DFGIESKRRF 50	LVAS 60
	70	80	90	100		
m127.pep	RNITLLLVLFSLAFIW.	SAOTOTTAL.	SMEAUAAAIRRA	100 ATKELIMCLS	- 110 GSTĪRSĀTOO	120
a127						
a171	RNITLLLVLFSLAFIW:	PUĞTĞIPUL	SHI AVAAAVVV	ATKELIMCLS	GSILRSATOO	YSVG
	70	80	90	100	110	120

m127.pep	130 DYIEINGLRGRVVD	140 TNI.I.NTI.MM(	150 DVGPNPLVGOL	160 AGTTVSEPNS	170	180
miz/.pcp		1		111111111		1111111
a127	DYIEINGLRGRVVD	INLLNTLMM	QVGPNPLVGQL	AGTTVSFPNS	LLLSHPVRR	DNILGDY
	130	140	150	160	170	180
	190	200	210	220	230	240
m127.pep	VIHTVEIPVPIHLD					
• •	111111111111111	111111111		111ÎL 1111	111111111	11:111
a127	VIHTVEIPVPIHLD	SDEAVCRLKA	AVLEPLCAPYI	PAIQRHLENV	'QAEKLFITP	AAKPRVT
	190	200	210	220	230	240
	250	260	270	280	290	
m127.pep	RVPYDDKAYRIIVR	FASPVSKRLE	EIQQAVMDEFL	RVQYRLLNHF	AGSETLX	
	1111111111111111	11111111		1111111:1	111111	
a127	RVPYDDKAYRIIVR	FASPVSKRLE	EIQQAVMDEFL	RVQYRLLNYF	AGSETLX	
	250	260	270	280	290	

# The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 481>: g128.seq

1	atgattgaca	acgCActgct	ccacttgggc	gaagaaccCC	GTTTTaatca
51	aatccaaacc	gaagACAtca	AACCCGCCGT	CCAAACCGCC	ATCGCCGAAG
101	CGCGCGGACA	AATCGCCGCC	GTCAAAGCGC	AAACGCACAC	CGGCTGGGCG
151	AACACCGTCG	AGCGTCTGAC	CGGCATCACC	GAACGCGTCG	GCAGGATTTG
201	GGGCGTCGTG	TCCCATCTCA	ACTCCGTCGT	CGACACGCCC	GAACTGCGCG
251	CCGTCTATAA	CGAACTGATG	CCTGAAATCA	CCGTCTTCTT	CACCGAAATC
301	GGACAAGACA	TCGAACTGTA	CAACCGCTTC	AAAACCATCA	AAAATTCCCC
351		ACGCTTTCCC			
401	TGCGCGATTT	CGTATTGAGC	GGCGCGGAAC	TGCCGCCCGA	ACGGCAGGCA
451		AACTGCAAAC			
501	CCAAAACGTC	CTAGACGCGA	CCGACGCGTT	CGGCATTTAC	TTTGACGATG
551		TGCCGGCATT			
601		GCGAAGGCAA			
651		GCCGTTATCC			
701		CGCCTACGTT			
751		ACACCGCCAA			
801		CTGCTCGGCT			
851		GGACACGCCC			
901		CCAAACCCTA			
951		GAACACCTCG			
1001		CGAAAAACTG			
1051		AATACTTCCC			
1101		AAACTCTACG			
1151		AGACGTGCGC			
1201		TTTATATGGA			
1251		AACGACtaca			
1301		CACCGCCTAC			
1351		CGCGTTTAAG			
1401		GGACTGCACC			
1451		CAAcggcgtA			
1501		ACTTCGTTTG			
1551		AccgGCGAGC			
1601		AAACTTCCAG			
1651		TCGATATGAT			
1701		CAGCAGGTTT			
1751		CGAATACAAC			
1801		ATTCCGCAGG			
1851	CAGCACCGAT	GCCTACGCCG	CCTTTGAAGA	AAGcGACGac	gtcGCCGCCA
1901	CAGGCAAACG	CTTCTGGCAA	GAAAtccttg	ccgtcggcgg	ctCCCGCAGC
1951	gcgGCGGAAT	CCTTCAAAGC	CTTCCGCGGA	CGCGAACCGA	GCATAGACGC

WO 99/57280



```
2001 ACTGCTGCGC CAmageggtT TCGACAACGC gGCttgA
  This corresponds to the amino acid sequence <SEQ ID 482; ORF 128.ng>:
       g128.pep
                 MIDNALLHLG EEPRFNQIQT EDIKPAVQTA IAEARGQIAA VKAQTHTGWA
                NTVERLTGIT ERVGRIWGVV SHLNSVVDTP ELRAVYNELM PEITVFFTEI
             51
                 GQDIELYNRF KTIKNSPEFA TLSPAQKTKL DHDLRDFVLS GAELPPERQA
            101
                 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
                 AAQSEGKTGY KIGLQIPHYL AVIQYAGNRE LREQIYRAYV TRASELSNDG
            201
                KFDNTANIDR TLENALKTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
            251
           301 ARRAKPYAEK DLAEVKAFAR EHLGLADPQP WDLSYAGEKL REAKYAFSET
           351 EVKKYFPVGK VLAGLFAQIK KLYGIGFAEK TVPVWHKDVR YFELQQNGKT
           401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFA DGTLQLPTAY LVCNFAPPVG
           451 GKEARLSHDE ILTLFHETGH GLHHLLTQVD ELGVSGINGV EWDAVELPSQ
           501 FMENFVWEYN VLAQMSAHEE TGEPLPKELF DKMLAAKNFQ RGMFLVRQME
           551 FALFDMMIYS ESDECRLKNW QQVLDSVRKE VAVIQPPEYN RFANSFGHIF
           601 AGGYSAGYYS YAWAEVLSTD AYAAFEESDD VAATGKRFWQ EILAVGGSRS
           651 AAESFKAFRG REPSIDALLR QSGFDNAA*
 The following partial DNA sequence was identified in N. meningitidis <SEQ ID 483>:
      m128.seq
                (partial)
                ATGACTGACA ACGCACTGCT CCATTTGGGC GAAGAACCCC GTTTTGATCA
             1
                AATCAAAACC GAAGACATCA AACCCGCCCT GCAAACCGCC ATCGCCGAAG
            51
                CGCGCGAACA AATCGCCGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA
           101
                AACACTGTCG AACCCCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
                GGGCGTGGTG TCGCACCTCA ACTGCGTCGC CGACACGCCC GAACTGCGCG
           201
           251 CCGTCTATAA CGAACTGATG CCCGAAATCA CCGTCTTCTT CACCGAAATC
                GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
           301
           351 CGAATTCGAC ACCCTCTCCC CCGCACAAAA AACCAAACTC AACCAC
               TACGCCAGCG AAAAACTGCG CGAAGCCAAA TACGCGTTCA GCGAAACCGA
            51 WGTCAAAAAA TAYTTCCCYG TCGGCAAWGT ATTAAACGGA CTGTTCGCCC
           101 AAMTCAAAAA ACTMTACGGC ATCGGATTTA CCGAAAAAAC YGTCCCCGTC
           151 TGGCACAAAG ACGTGCGCTA TTKTGAATTG CAACAAAACG GCGAAMCCAT
           201 AGGCGGCGTT TATATGGATT TGTACGCACG CGAAGGCAAA CGCGGCGGCG
               CGTGGATGAA CGACTACAAA GGCCGCCGCC GTTTTTCAGA CGGCACGCTG
               CAAYTGCCCA CCGCCTACCT CGTCTGCAAC TTCGCCCCAC CCGTCGGCGG
               CAGGGAAGCC CGCYTGAGCC ACGACGAAAT CCTCATCCTC TTCCACGAAA
          351
          401 CCGGACACGG GCTGCACCAC CTGCTTACCC AAGTGGACGA ACTGGGCGTA
               TCCGGCATCA ACGGCGTAKA ATGGGACGCG GTCGAACTGC CCAGCCAGTT
          451
               TATGGAAAAT TTCGTTTGGG AATACAATGT CTTGGCACAA mTGTCAGCCC
          551 ACGAAGAAAC CGGCGTTCCC yTGCCGAAAG AACTCTTsGA CAAAwTGCTC
          601 GCCGCCAAAA ACTTCCAASG CGGCATGTTC yTsGTCCGGC AAWTGGAGTT
          651 CGCCCTCTTT GATATGATGA TTTACAGCGA AGACGACGAA GGCCGTCTGA
          701 AAAACTGGCA ACAGGTTTTA GACAGCGTGC GCAAAAAAGT CGCCGTCATC
               CAGCCGCCCG AATACAACCG CTTCGCCTTG AGCTTCGGCC ACATCTTCGC
               AGGCGGCTAT TCCGCAGCTn ATTACAGCTA CGCGTGGGCG GAAGTATTGA
               GCGCGGACGC ATACGCCGCC TTTGAAGAAA GCGACGATGT CGCCGCCACA
          851
          901 GGCAAACGCT TTTGGCAGGA AATCCTCGCC GTCGGGGAAT CGCGCAGCGG
              NGCAGAATCC TTCAAAGCCT TCCGCGGCCG CGAACCGAGC ATAGACGCAC
              TCTTGCGCCA CAGCGGTTTC GACAACGCGG TCTGA
This corresponds to the amino acid sequence <SEQ ID 484; ORF 128>:
     m128.pep
                (partial)
               MTDNALLHLG EEPRFDQIKT EDIKPALQTA IAEAREQIAA IKAQTHTGWA
           51 NTVEPLTGIT ERVGRIWGVV SHLNCVADTP ELRAVYNELM PEITVFFTEI
          101 GQDIELYNRF KTIKNSPEFD TLSPAQKTKL NH
     //
              YASEKLREAK YAFSETXVKK YFPVGXVLNG LFAQXKKLYG IGFTEKTVPV
           51 WHKDVRYXEL QQNGEXIGGV YMDLYAREGK RGGAWMINDYK GRRRFSDGTL
          101 QLPTAYLVCN FAPPVGGREA RLSHDEILIL FHETGHGLHH LLTQVDELGV
          151 SGINGVXWDA VELPSQFMEN FVWEYNVLAQ XSAHEETGVP LPKELXDKXL
              AAKNFOXGMF XVRQXEFALF DMMIYSEDDE GRLKNWQQVL DSVRKKVAVI
              QPPEYNRFAL SFGHIFAGGY SAAXYSYAWA EVLSADAYAA FEESDDVAAT
```

301 GKRFWQEILA VGXSRSGAES FKAFRGREPS IDALLRHSGF DNAV* Computer analysis of this amino acid sequence gave the following results: Homology with a predicted ORF from N. gonorrhoeae

ORF 128 shows 91.7% identity over a 475 aa overlap with a predicted ORF (ORF 128.ng) from N. gonorrhoeae:

m128/g128

g128.pep	10 20 30 40 50 60 MIDNALLHLGEEPRFNQIQTEDIKPAVQTAIAEARGQIAAVKAQTHTGWANTVERLTGIT
g128.pep m128	70 80 90 100 110 120 ERVGRIWGVVSHLNSVVDTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFA
g128.pep m128	130 140 150 160 170 180 TLSPAQKTKLDHDLRDFVLSGAELPPERQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY           TLSPAQKTKLNH 130
g128.pep m128	// 340 350 360 YAGEKLREAKYAFSETEVKKYFPVGKVLAG   :
g128.pep	370 380 390 400 410 420  LFAQIKKLYGIGFAEKTVPVWHKDVRYFELQQNGKTIGGVYMDLYAREGKRGGAWMNDYK
g128.pep m128	430 440 450 460 470 480 GRRRFADGTLQLPTAYLVCNFAPPVGGKEARLSHDEILTLFHETGHGLHHLLTQVDELGV     :
g128.pep m128	490 500 510 520 530 540  SGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGEPLPKELFDKMLAAKNFQRGMF
g128.pep	550 560 570 580 590 600  LVRQMEFALFDMMIYSESDECRLKNWQQVLDSVRKEVAVIQPPEYNRFANSFGHIFAGGY
g128.pep	610 620 630 640 650 660  SAGYYSYAWAEVLSTDAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKAFRGREPS   :         :

BNSDOCID: <WO___9957280A2_l_>

 $s = \sup_{t \in \mathcal{S}_{t}} |s_{t}(t)| \leq s + \sup_{t \in \mathcal{S}_{t}} |s_{t}(t)|$ 

670 679
g128.pep IDALLRQSGFDNAAX
|||||:||||:
m128 IDALLRHSGFDNAVX
340

### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 485>:

```
a128.seg
          ATGACTGACA ACGCACTGCT CCATTTGGGC GAAGAACCCC GTTTTGATCA
      1
         AATCAAAACC GAAGACATCA AACCCGCCCT GCAAACCGCC ATTGCCGAAG
      51
     101
         CGCGCGAACA AATCGCCGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA
         AACACTGTCG AACCCCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
    201
         GGGCGTGGTG TCGCACCTCA ACTCCGTCAC CGACACGCCC GAACTGCGCG
    251
         CCGCCTACAA TGAATTAATG CCCGAAATTA CCGTCTTCTT CACCGAAATC
         GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAACTCCCC
    301
         CGAGTTCGAC ACCCTCTCCC ACGCGCAAAA AACCAAACTC AACCACGATC
    351
    401
         TGCGCGATTT CGTCCTCAGC GGCGCGGAAC TGCCGCCCGA ACAGCAGGCA
         GAATTGGCAA AACTGCAAAC CGAAGGCGCG CAACTTTCCG CCAAATTCTC
    451
         CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTTAC TTTGACGATG
    501
         CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCT
    551
         GCCGCGCAAA GCGAAGGCAA AACAGGCTAC AAAATCGGTT TGCAGATTCC
    601
    651
         GCACTACCTC GCCGTCATCC AATACGCCGA CAACCGCAAA CTGCGCGAAC
         AAATCTACCG CGCCTACGTT ACCCGCGCCA GCGAGCTTTC AGACGACGGC
    701
         AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCCCTGCA
    751
    801
         AACCGCCAAA CTGCTCGGCT TCAAAAACTA CGCCGAATTG TCGCTGGCAA
         CCAAAATGGC GGACACCCC GAACAAGTTT TAAACTTCCT GCACGACCTC
         GCCCGCCGC CCAAACCCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
    901
         CTTCGCCCGC GAAAGCCTCG GCCTCGCCGA TTTGCAACCG TGGGACTTGG
    951
         GCTACGCCGG CGAAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
   1001
   1051
         GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTATTAAACG GACTGTTCGC
         CCAAATCAAA AAACTCTACG GCATCGGATT TACCGAAAAA ACCGTCCCCG
   1151
         TCTGGCACAA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCGAAACC
   1201
         ATAGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
         CGCGTGGATG AACGACTACA AAGGCCGCCG CCGTTTTTCA GACGGCACGC
   1251
         TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCACCCC GCCCGTCGGC
   1301
         GGCAAAGAAG CCCGCTTGAG CCATGACGAA ATCCTCACCC TCTTCCACGA
   1351
         AACCGGACAC GGCCTGCACC ACCTGCTTAC CCAAGTCGAC GAACTGGGCG
   1451
         TATCCGGCAT CAACGGCGTA GAATGGGACG CAGTCGAACT GCCCAGTCAG
         TTTATGGAAA ATTTCGTTTG GGAATACAAT GTCTTGGCGC AAATGTCCGC
   1501
   1551
         CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAACTCTTC GACAAAATGC
   1601
         TCGCCGCCAA AAACTTCCAA CGCGGAATGT TCCTCGTCCG CCAAATGGAG
         TTCGCCCTCT TTGATATGAT GATTTACAGC GAAGACGACG AAGGCCGTCT
   1701
         GAAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAGAA GTCGCCGTCG
   1751
         TCCGACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCACATCTTC
         GCAGGCGGCT ATTCCGCAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
   1801
   1851
         GAGCGCGGAC GCATACGCCG CCTTTGAAGA AAGCGACGAT GTCGCCGCCA
         CAGGCAAACG CTTTTGGCAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
         GCGGCAGAAT CCTTCAAAGC CTTCCGCGGA CGCGAACCGA GCATAGACGC
   1951
        ACTCTTGCGC CACAGCGGCT TCGACAACGC GGCTTGA
```

## This corresponds to the amino acid sequence <SEQ ID 486; ORF 128.a>:

-		· · · · · · · · · · · · · · · · · · ·	4 1	- 100, OIG	120.a.
a128.pep				•	
1	MTDNALLHLG	EEPRFDQIKT	EDIKPALQTA	IAEAREQIAA	IKAOTHTGWA
51	NTVEPLTGIT	ERVGRIWGVV	SHLNSVTDTP	ELRAAYNELM	PEITVFFTEI
101	GQDIELYNRF	KTIKNSPEFD	TLSHAQKTKL	NHDLRDFVLS	GAELPPEOOA
151	ELAKLQTEGA	QLSAKFSQNV	LDATDAFGIY	FDDAAPLAGI	PEDALAMFAA
201	AAQSEGKTGY	KIGLQIPHYL	AVIQYADNRK	LREOIYRAYV	TRASELSDDG
251	KFDNTANIDR	TLENALQTAK	LLGFKNYAEL	SLATKMADTP	EOVLNFLHDI.
301	ARRAKPYAEK	DLAEVKAFAR	ESLGLADLQP	WDLGYAGEKL	REAKYAFSET
351	EVKKYFPVGK	VLNGLFAQIK	KLYGIGFTEK	TVPVWHKDVR	YELLOONGET
401	IGGVYMDLYA	REGKRGGAWM	NDYKGRRRFS	DGTLOLPTAY	LVCNFTPPVG
451	GKEARLSHDE	ILTLFHETGH	GLHHLLTQVD	ELGVSGINGV	EWDAVEL PSO
501	FMENFVWEYN	VLAQMSAHEE	TGVPLPKELF	DKMLAAKNFQ	RGMFLVROME
				• • • • • • • • • • • • • • • • • • • •	-

FALFDMMIYS EDDEGRLKNW QQVLDSVRKE VAVVRPPEYN RFANSFGHIF AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRFWQ EILAVGGSRS

AAESFKAFRG REPSIDALLR HSGFDNAA* m128/a128 66.0% identity in 677 aa overlap MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIKAQTHTGWANTVEPLTGIT m128.pep a128 MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIKAQTHTGWANTVEPLTGIT ERVGRIWGVVSHLNCVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD m128.pep ERVGRIWGVVSHLNSVTDTPELRAAYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD a128 TLSPAOKTKLNH-m128.pep 111 1111111 TLSHAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY a128 m128.pep FDDAAPLAGIPEDALAMFAAAAQSEGKTGYKIGLQIPHYLAVIQYADNRKLREQIYRAYV a128 m128.pep a128 TRASELSDDGKFDNTANIDRTLENALQTAKLLGFKNYAELSLATKMADTPEOVLNFLHDL -----YASEKLREAKYAFSETXVKKYFPVGX m128.pep a128 ARRAKPYAEKDLAEVKAFARESLGLADLOPWDLGYAGEKLREAKYAFSETEVKKYFPVGK VLNGLFAQXKKLYGIGFTEKTVPVWHKDVRYXELQQNGEXIGGVYMDLYAREGKRGGAWM m128.pep a128 VLNGLFAQIKKLYGIGFTEKTVPVWHKDVRYFELQQNGETIGGVYMDLYAREGKRGGAWM

ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ a128 XGMFXVRQXEFALFDMMIYSEDDEGRLKNWQQVLDSVRKKVAVIQPPEYNRFALSFGHIF m128.pep RGMFLVRQMEFALFDMMIYSEDDEGRLKNWQQVLDSVRKEVAVVRPPEYNRFANSFGHIF a128 

NDYKGRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILILFHETGHGLHHLLTQVD

ELGVSGINGVXWDAVELPSQFMENFVWEYNVLAQXSAHEETGVPLPKELXDKXLAAKNFQ

m128.pep

m128.pep

a128

m128.pep	400	410	420	430	440	450	
milo.bcb	AGGISA	AMAISIAWA	FATZADAĀĀĀ	EESDDVAATG	KRFWQEILAV	GXSRSGAE.	SEKAFRG
a128			1   1   1   1   1   1   1   1	  EESDDVAATG			
		610	620	630	VKEMÕETTAA		SFKAFRG
			020	630	640	650	660
	460	470					
m128.pep	REPSID	ALLRHSGF	XVANC				
		11111111					
a128	REPSID	ALLRHSGF	NAAX				
		670					

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 487>:

```
1
       ATGATTGACA ACGCACTGCT CCACTTGGGC GAAGAACCCC GTTTTAATCA
   51 AATCAAAACC GAAGACATCA AACCCGCCGT CCAAACCGCC ATCGCCGAAG
  101 CGCGCGGACA AATCGCCGCC GTCAAAGCGC AAACGCACAC CGGCTGGGCG
  151 AACACCGTCG AGCGTCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
  201 GGGCGTCGTG TCCCATCTCA ACTCCGTCGT CGACACGCCC GAACTGCGCG
  251 CCGTCTATAA CGAACTGATG CCTGAAATCA CCGTCTTCTT CACCGAAATC
      GGACAAGACA TCGAACTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
  301
      CGAATTTGCA ACGCTTTCCC CCGCACAAAA AACCAAGCTC GATCACGACC
  351
      TGCGCGATTT CGTATTGAGC GGCGCGGAAC TGCCGCCCGA ACGGCAGGCA
  451
      GAACTGGCAA AACTGCAAAC CGAAGGCGCG CAACTTTCCG CCAAATTCTC
      CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTTAC TTTGACGATG
  501
 551 CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCC
 601 GCCGCGCAAA GCGAAGGCAA AACAGGTTAC AAAATCGGCT TGCAGATTCC
      GCACTACCTT GCCGTTATCC AATACGCCGG CAACCGCGAA CTGCGCGAAC
  651
 701 AAATCTACCG CGCCTACGTT ACCCGTGCCA GCGAACTTTC AAACGACGGC
 751
      AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCATTGAA
      AACCGCCAAA CTGCTCGGCT TTAAAAATTA CGCCGAATTG TCGCTGGCAA
 801
 851 CCAAAATGGC GGACACGCCC GAACAGGTTT TAAACTTCCT GCACGACCTC
 901 GCCCGCCGCG CCAAACCCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
 951 CTTCGCCCGC GAACACCTCG GTCTCGCCGA CCCGCAGCCG TGGGACTTGA
1001 GCTACGCCGG CGAAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTTCTGGCAG GCCTGTTCGC
1101 CCAAATCAAA AAACTCTACG GCATCGGATT CGCCGAAAAA ACCGTTCCCG
1151 TCTGGCACAA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCAAAACC
1201 ATCGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
1251 CGCGTGGATG AACGACTACA AAGGCCGCCG CCGCTTTGCC GACGGCACGC
      TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCGCCCC GCCCGTCGGC
1301
1351 GGCAAAGAAG CGCGTTTAAG CCACGACGAA ATCCTCACCC TCTTCCACGA
1401 AACCGGCCAC GGACTGCACC ACCTGCTTAC CCAAGTGGAC GAACTGGGCG
1451 TGTCCGGCAT CAACGGCGTA AAA
```

This corresponds to the amino acid sequence <SEQ ID 488; ORF 128-1.ng>:

```
1 MIDNALLHLG EEPRFNQIKT EDIKPAVQTA IAEARGQIAA VKAQTHTGWA
51 NTVERLTGIT ERVGRIWGVV SHLNSVVDTP ELRAVYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFA TLSPAQKTKL DHDLRDFVLS GAELPPERQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSEGKTGY KIGLQIPHYL AVIQYAGRRE LREQIYRAYV TRASELSNDG
251 KFDNTANIDR TLENALKTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
301 ARRAKPYAEK DLAEVKAFAR EHLGLADPQP WDLSYAGEKL REAKYAFSET
351 EVKKYFPVGK VLAGLFAQIK KLYGIGFAEK TVPVWHKDVR YFELQQNGKT
401 IGGVYMDLYA REGKRGGAWM NDYKGRRFFA DGTLQLPTAY LVCNFAPPVG
451 GKEARLSHDE ILTLFHETGH GLHHLLTQVD ELGVSGINGV K
```

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 489>:

```
1 ATGACTGACA ACGCACTGCT CCATTTGGGC GAAGAACCCC GTTTTGATCA
51 AATCAAAACC GAAGACATCA AACCCGCCCT GCAAACCGCC ATCGCCGAAG
101 CGCGCGAACA AATCGCCGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA
151 AACACTGTCG AACCCCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
```

```
201 GGGCGTGGTG TCGCACCTCA ACTCCGTCGC CGACACGCCC GAACTGCGCG
     CCGTCTATAA CGAACTGATG CCCGAAATCA CCGTCTTCTT CACCGAAATC
 301 GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
 351 CGAATTCGAC ACCCTCTCCC CCGCACAAAA AACCAAACTC AACCACGATC
      TGCGCGATTT CGTCCTCAGC GGCGCGGAAC TGCCGCCCGA ACAGCAGGCA
 451 GAACTGCAA AACTGCAAAC CGAAGGCGCG CAACTTTCCG CCAAATTCTC
 501 CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTTAC TTTGACGATG
 551
      CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCC
 601 GCCGCGCAAA GCGAAAGCAA AACAGGCTAC AAAATCGGCT TGCAGATTCC
 651 ACACTACCTC GCCGTCATCC AATACGCCGA CAACCGCGAA CTGCGCGAAC
 701 AAATCTACCG CGCCTACGTT ACCCGCGCCA GCGAACTTTC AGACGACGGC
      AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGCAA ACGCCCTGCA
 801 AACCGCCAAA CTGCTCGGCT TCAAAAACTA CGCCGAATTG TCGCTGGCAA
 851 CCAAAATGGC GGACACGCCC GAACAAGTTT TAAACTTCCT GCACGACCTC
 901 GCCCGCCGCG CCAAACCCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
 951 CTTCGCCCGC GAAAGCCTGA ACCTCGCCGA TTTGCAACCG TGGGACTTGG
1001 GCTACGCCAG CGAAAAACTG CGCGAAGCCA AATACGCGTT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTATTAAACG GACTGTTCGC
      CCAAATCAAA AAACTCTACG GCATCGGATT TACCGAAAAA ACCGTCCCCG
1151
     TCTGGCACAA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCGAAACC
1201 ATAGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
1251
      CGCGTGGATG AACGACTACA AAGGCCGCCG CCGTTTTTCA GACGGCACGC
1301 TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCGCCCC ACCCGTCGGC
1351 GGCAGGGAAG CCCGCCTGAG CCACGACGAA ATCCTCATCC TCTTCCACGA
1401 AACCGGACAC GGGCTGCACC ACCTGCTTAC CCAAGTGGAC GAACTGGGCG
1451
     TATCCGGCAT CAACGCCGTA GAATGGGACG CGGTCGAACT GCCCAGCCAG
1501 TTTATGGAAA ATTTCGTTTG GGAATACAAT GTCTTGGCAC AAATGTCAGC
1551 CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAACTCTTC GACAAAATGC
     TCGCCGCCAA AAACTTCCAA CGCGGCATGT TCCTCGTCCG GCAAATGGAG
1651 TTCGCCCTCT TTGATATGAT GATTTACAGC GAAGACGACG AAGGCCGTCT
1701 GAAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAAAA GTCGCCGTCA
1751
     TCCAGCCGCC CGAATACAAC CGCTTCGCCT TGAGCTTCGG CCACATCTTC
1801 GCAGGCGGCT ATTCCGCAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
1851 GAGCGCGGAC GCATACGCCG CCTTTGAAGA AAGCGACGAT GTCGCCGCCA
1901 CAGGCAAACG CTTTTGGCAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
1951 GCGGCAGAAT CCTTCAAAGC CTTCCGCGGC CGCGAACCGA GCATAGACGC
2001 ACTCTTGCGC CACAGCGGTT TCGACAACGC GGTCTGA
```

## This corresponds to the amino acid sequence <SEQ ID 490; ORF 128-1>:

```
1 MTDNALLHLG EEPRFDQIKT EDIKPALQTA IAEAREQIAA IKAQTHTGWA
51 NTVEPLTGIT ERVGRIWGVV SHLNSVADTP ELRAVYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFD TLSPAQKTKL NHDLRDFVLS GAELPPEQQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSESKTGY KIGLQIPHYL AVIQYADNRE LREQIYRAYV TRASELSDDG
251 KFDNTANIDR TLANALQTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
301 ARRAKPYAEK DLAEVKAFAR ESLNLADLQP WDLGYASEKL REAKYAFSET
351
    EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQONGET
401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTLQLPTAY LVCNFAPPVG
    GREARLSHDE ILILFHETGH GLHHLLTQVD ELGVSGINGV EWDAVELPSO
    FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME
501
551
    FALFDMMIYS EDDEGRLKNW QQVLDSVRKK VAVIQPPEYN RFALSFGHIF
601 AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRFWQ EILAVGGSRS
651 AAESFKAFRG REPSIDALLR HSGFDNAV*
```

#### ml28-1/g128-1 94.5% identity in 491 aa overlap

	10	20	30	40	50	60
g128-1.pep	MIDNALLHLGEEPR					
		1:111111	11:111111	1	1111111111	11111
m128-1	MTDNALLHLGEEPR	FDQIKTEDIK	PALQTAIAEA	REQIAAIKAQ	THTGWANTVE	PLTGIT
	10	20	30	40	50	60
	70	80	90	100	110	120
g128-1.pep	ERVGRIWGVVSHLN					
		11:11111	1111111111	1111111111	1111111111	11111
m128-1	ERVGRIWGVVSHLN					
	70	80	90	100	110	120

~100.1	130	140	150	160	170	180
g128-1.pep	TLSPAQKTKLDHD	LRDFVLSGAE	LPPERQAELA	KLQTEGAQLS		
m128-1	1   1   1   1   1   1   1   1   1   1				1 1 1 2 3 3 4 4 4	
	TLSPAQKTKLNHD	LKDFVLSGAE 140	TPPEQQAELA	KLQTEGAQLS	AKFSQNVLD	ATDAFGIY
	130	140	150	160	170	180
	190	200	210	220		
g128-1.pep	FDDAAPLAGIPED	ALAMFAAAAO	SECKTOVKIC	220 LOT DUVI AUT	230	240
	1 1 1 1 1 1 1 1 1 1 1 1 1	1	112111111	4 1 1 <b>1</b> 1 1 1 1 1 1 1 1		
m128-1	FDDAAPLAGIPED	ALAMFAAAAQ	SESKTGYKIG	LOIPHYLAVT	IIIIIIIII Tyanadatee	OTYPARK
	190	200	210	220	230	240
						240
g128-1.pep	250	260	270	280	290	300
3-20 1.pcp	TRASELSNDGKFDN	NTANIDRTLEI 	NALKTAKLLGI	KNYAELSLAT	KMADTPEQV	LNFLHDL
m128-1		!                 JTANTDOTIAN	:	1111111111	11111111	1111111
	TRASELSDDGKFDN 250	260	270	KNYAELSLAT 280		
			270	. 280	290	300
	310	320	330	340	350	
g128-1.pep	ARRAKPYAEKDLAE	VKAFAREHLO	LADPOPWDLS	YAGEKLREAK	JOU Vàrgemetry:	360
100 1		11111111111		11.11111	111111111	
m128-1	ARRAKPYAEKDLAE	VKAFARESLN	LADLQPWDLG	YASEKLREAK	YAFSETEVKI	KYFPVGK
	310	320	330	340	350	360
	370	380	200			
g128-1.pep	VLAGLFAQIKKLYG	VOUTTYTATATATAT	390	400	410	420
_	VINGLEROIKKIVO	111:11111	HILLIIII	JONGKTIGGV	YMDLYAREGI	CRGGAWM
m128-1	VLNGLFAQIKKLYG	IGFTEKTVPV	WHKDVRYFEI	)	MADI YARRA	
	370	380	390	400	410	AGGAWM 420
					110	420
g128-1.pep	430	440	450	460	470	480
g126-1.pep	NDYKGRRRFADGTLO	OLPTAYLVCN	Fappvggkeai	LSHDEILTLE	HETGHGLHH	LLTQVD
m128-1	_		1			
	NDYKGRRRFSDGTLC	440	PAPPVGGREAR	LSHDEILILE		LLTQVD
	100	440	450	460	470	480
	490					
g128-1.pep	ELGVSGINGVK					
m128-1	ELGVSGINGVEWDAV	ELPSQFMENF	<b>WEYNVLAQM</b>	SAHEETGVPL	PKELFDKMT.:	AAKMEO
	490	500	510	520	530	540
						5.10

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 491>:

1 ATGACTGACA ACGCACTGCT CCATTTGGGC GAAGAACCCC GTTTTGATCA

1	איריים איריים ע				
51	AIGACIGAC	A ACGCACTGC	r ccatttggg	GAAGAACCCC	GTTTTGATCA
	AATCAAAAC	C GAAGACATC	AACCCGCCC	C GCAAACCGCC	TOTOGGGATA
101	CGCGCGAAC	A AATCGCCGC	ATCAAAGCC	AAACGCACAC	CCCCCCCCC
151	MACACTGTC	G AACCCCTGAC	CGGCATCAC	GAACGCGTCC	CCACCAMmm
201	GGGCGTGGT	F TCGCACCTCF	ACTCCGTCAC	CGACACCCC	CARCINGGGGG
251	CCGCCTACA	A TGAATTAATC	CCCGAAATT	CCGTCTTCTT	CACCCAAA
301	GGACAAGACA	1 TCGAGCTGTA	CAACCGCTTC	באסטרכיטייטא באם "	7 7 7 7 CMCCCC
351	CGAGTTCGAC	ACCCTCTCCC	ACGCGCAAAA	AACCAAACTC	**********
401	I GCGCGATTT	CGTCCTCAGC	GGCGCGGAAC	TGCCGCCCGA	ACACCACCCA
451	GAATTGGCAA	AACTGCAAAC	CGAAGGCGCG	CAACTTTCCC	CCARAGGCA
501	CCAAAACGTC	CTAGACGCGA	CCGACGCGTT	CCCCATTCCC	CCAAATTCTC
551	CCGCACCGCT	TGCCGGCATT	CCCGAAGACG	CCCTCCCCA	TTTGACGATG
601	GCCGCGCAAA	GCGAAGGCAA	AACAGGCTAC	AAAATGCCAT	GTTTGCCGCT
651	GCACTACCTC	GCCGTCATCC	AATACGCCGA	CAACCCCAA	TGCAGATTCC
701	AAATCTACCG	CGCCTACGTT	ACCCGCGCGA	CCCACCGCAAA	CIGCGCGAAC
751	AAATTCGACA	ACACCGCCAA	CATCGACCCC	ACCOMOCTITO	AGACGACGGC
801	AACCGCCAAA	CTGCTCGGCT	TCAAAAACCGC	ACGCTCGAAA	ACGCCCTGCA
851	CCAAAATGGC	GGACACCCCC	CONTANACIA	CGCCGAATTG	TCGCTGGCAA
901	GCCCGCCGCG	GGACACCCCC	GAACAAGTTT	TAAACTTCCT	GCACGACCTC
951	CTTCGCCCCC	CCAAACCCTA GAAAGCCTCG	CGCCGAAAAA	GACCTCGCCG	AAGTCAAAGC
1001		CGAAAAACTG		TTTGCAACCG	TGGGACTTGG
1051		AATACTTCCC		AATACGCATT	CAGCGAAACC
1101	CCNNATCANA	AATACTTCCC	CGTCGGCAAA	GTATTAAACG	GACTGTTCGC
1151	TCTCCCO	AAACTCTACG	GCATCGGATT	TACCGAAAAA	ACCGTCCCCG
1201	1C1GGCACAA	AGACGTGCGC	TATTTTGAAT	TCCAACAAAA	CCCCCCAAAA
TZ () I	WINGCCGCC	TTTATATGGA	TTTGTACGCA	CGCGAAGGCA	AACCCCCCCC

```
1251 CGCGTGGATG AACGACTACA AAGGCCGCCG CCGTTTTTCA GACGCACGC
1301 TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCACCCC GCCCGTCGGC
1351 GGCAAAGAAG CCCGCTTGAG CCATGACGAA ATCCTCACCC TCTTCCACGA
1401 AACCGGACAC GGCCTGCACC ACCTGCTTAC CCAAGTCGAC GAACTGGGCG
1451 TATCCGGCAT CAACGGCGTA GAATGGGACG CAGTCGAACT GCCCAGTCAG
1501 TTTATGGAAA ATTTCGTTTG GGAATACAAT GTCTTGGCGC AAATGTCCGC
1551 CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAACTCTTC GACAAAATGC
1601 TCGCCGCCAA AAACTTCCAA CGCGGAATGT TCCTCGTCCG CCAAATGGAG
1651 TTCGCCCTCT TTGATATGAT GATTTACAGC GAAGACGACG AAGGCCGTCT
1701 GAAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAGAA GTCGCCGTCG
1751 TCCGACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCACATCTTC
1801 GCAGGCGGCT ATTCCGCAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
1851 GAGCGCGGAC GCATACGCCG CCTTTGAAGA AAGCGACGAT GTCGCCGCCA
1901 CAGGCAAACG CTTTTGGCAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
1951 GCGGCAGAAT CCTTCAAAGC CTTCCGCGGA CGCGAACCGA GCATAGACGC
2001 ACTCTTGCGC CACAGCGGCT TCGACAACGC GGCTTGA
```

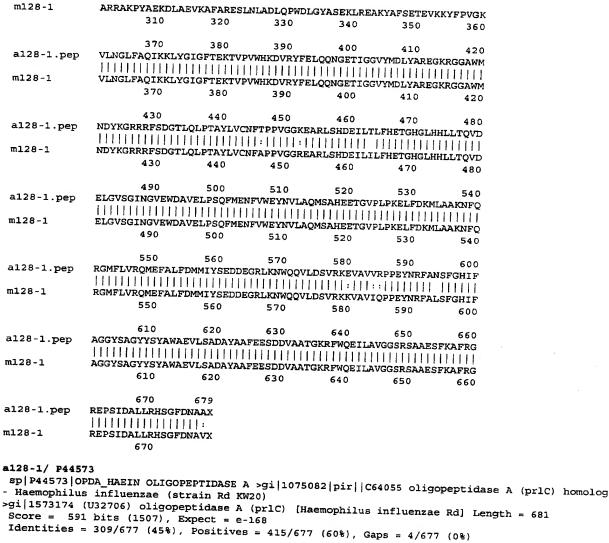
This corresponds to the amino acid sequence <SEQ ID 492; ORF 128-1.a>: a128-1.pep

```
1 MTDNALLHLG EEPRFDQIKT EDIKPALQTA IAEAREQIAA IKAQTHTGWA
 51 NTVEPLTGIT ERVGRIWGVV SHLNSVTDTP ELRAAYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFD TLSHAQKTKL NHDLRDFVLS GAELPPEQQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSEGKTGY KIGLQIPHYL AVIQYADNRK LREQIYRAYV TRASELSDDG
251 KFDNTANIDR TLENALQTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
301 ARRAKPYAEK DLAEVKAFAR ESLGLADLQP WDLGYAGEKL REAKYAFSET
351 EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQQNGET
401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTLQLPTAY LVCNFTPPVG
451 GKEARLSHDE ILTLFHETGH GLHHLLTQVD ELGVSGINGV EWDAVELPSQ
501 FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVROME
551 FALFDMMIYS EDDEGRLKNW QQVLDSVRKE VAVVRPPEYN RFANSFGHIF
601 AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRFWQ EILAVGGSRS
651 AAESFKAFRG REPSIDALLR HSGFDNAA*
```

m128-1/a128-1 97.8% identity in 677 aa overlap

al28-1.pep ml28-1	10 MTDNALLHLGEEPR            MTDNALLHLGEEPR 10	ПІННІ	ШІШШ	40 AREQIAAIKAQ          AREQIAAIKAQ 40	ШШШ	
al28-1.pep m128-1	70 ERVGRIWGVVSHLN            ERVGRIWGVVSHLN 70	11:111111	:			
a128-1.pep	130 TLSHAQKTKLNHDI             TLSPAQKTKLNHDI 130	1111111111	ШШШ	ППППП	ШППП	
a128-1.pep	190 FDDAAPLAGIPEDA           FDDAAPLAGIPEDA 190	HIHHH	1:11111111	ĪШШПĪ	111111111	Ĩ111111
al28-1.pep	250 TRASELSDDGKFDN           TRASELSDDGKFDN 250	HHHHH	1111111111	1111111111	11111111	1111111
al28-1.pep	310 ARRAKPYAEKDLAE	320 VKAFARESLG	330 LADLQPWDLG	340 SYAGEKLREAK   :	350 YAFSETEVKI	360 KYFPVGK





- Haemophilus influenzae (strain Rd KW20)

Score = 591 bits (1507), Expect = e-168

- NALLHLGEEPRFDQIKTEDIKPALQTXXXXXXXXXXXXXXXXTHTGWANTVEPLTGITERV 63 Query: 4 P F QIK E I+PA++ H WN + PLT
- NPLLNIQGLPPFSQIKPEHIRPAVEKLIQDCRNTIEQVLKQPHFTWENFILPLTETNDRL 64 Sbjct: 5
- GRIWGVVSHLNSVTDTPELRAAYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFDTLS 123 Query: 64 R W VSHLNSV ++ ELR AY +P ++ + T +GQ LYN + +KNS EF
- Sbjct: 65 NRAWSPVSHLNSVKNSTELREAYQTCLPLLSEYSTWVGQHKGLYNAYLALKNSAEFADYS 124
- Query: 124 HAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIYFDD 183 AQK + + LRDF LSG L E+O
- ++ ++L+++FS NVLDAT + Sbjct: 125 IAQKKAIENSLRDFELSGIGLSEEKQQRYGEIVARLSELNSQFSNNVLDATMGWEKLIEN 184
- Query: 184 AAPLAGIPEDALAMFAAAAQSEGKTGYKIGLQIPHYLAVIQYADNRKLREQIYRAYVTRA 243
- A LAG+PE AL +A+S+G GY+ L+IP YL V+ Y +NR LRE++YRAY TRA Sbjct: 185 EAELAGLPESALQAAQQSAESKGLKGYRFTLEIPSYLPVMTYCENRALREEMYRAYATRA 244
- Query: 244 SELSDD-GKFDNTANIDRTLENALQTAKLLGFKNYAELSLATKMADTPEQVLNFLHDLAR 302 + GK+DN+ ++ L ++ AKLLGF Y ELSLATKMA+ P+QVL+FL LA SE
- Sbjct: 245 SEQGPNAGKWDNSKVMEEILTLRVELAKLLGFNTYTELSLATKMAENPQQVLDFLDHLAE 304
- Query: 303 RAKPYAEKDLAEVKAFARESLGLADLQPWDLGYAGEKLREAKYAFSETEVKKYFPVGKVL 362
- RAKP EK+L E+K + + G+ +L PWD+G+ EK ++ YA ++ E++ YFP +V+ Sbjct: 305 RAKPQGEKELQELKGYCEKEFGVTELAPWDIGFYSEKQKQHLYAINDEELRPYFPENRVI 364

. .

A HOLD BOOK STATE

```
Query: 363 NGLFAQIKKLYGIGFTE-KTVPVWHKDVRYFEL-QQNGETIGGVYMDLYAREGKRGGAWM 420
          +GLF IK+++ I E K V WHKDVR+F+L +N + G Y+DLYARE KRGGAWM
Sbjct: 365 SGLFELIKRIFNIRAVERKGVDTWHKDVRFFDLIDENDQLRGSFYLDLYAREHKRGGAWM 424
+D GR+R DG+++ P AYL CNF P+G K A +H+E+
Sbjct: 425 DDCIGRKRKLDGSIETPVAYLTCNFNAPIGNKPALFTHNEVTTLFHEFGHGIHHMLTQID 484
Query: 481 ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ 540
            V+GINGV WDAVELPSQFMEN+ WE LA +S H ETG PLPKE ++L AKNFQ
Sbjct: 485 VSDVAGINGVPWDAVELPSQFMENWCWEEEALAFISGHYETGEPLPKEKLTOLLKAKNFO 544
Query: 541 RGMFLVRQMEFALFDMMIYSEDDEGRLKNWQQVLDSVRKEVAVVRPPEYNRFANSFGHIF 600
           MF++RQ+EF +FD ++ D +
                                      L SV+ +VAV++ ++ R +SF HIF
Sbjct: 545 AAMFILRQLEFGIFDFRLHHTFDAEKTNQILDTLKSVKSQVAVIKGVDWARAPHSFSHIF 604
Query: 601 XXXXXXXXXXXXWAEVLSADAYAAFEESDDV-AATGKRFWQEILAVGGSRSAAESFKAFR 659
                    WAEVLSADAY+ FEE TGK F EIL GGS
Sbjct: 605 AGGYAAGYYSYLWAEVLSADAYSRFEEEGIFNPITGKSFLDEILTRGGSEEPMELFKRFR 664
Query: 660 GREPSIDALLRHSGFDN 676
         GREP +DALLRH G N
Sbjct: 665 GREPQLDALLRHKGIMN 681
The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 493>:
     g129.seq
           1 ATGCTTTCAC CTCCTCGGCG TAAAACGGCG GCACATCAAT CAAGCCGTCT
           51 TTCATTTGCG TGCGGAAAAA ATGCGGCGTG TTGCCGTGAT CAAAATCAAT
          101 ATCGTGCAGC ATCCAGCCCA AATCGCGGTT TGCCTCGCTT TCCGATAACG
          151 CCGACGCGG CAGCGGTTCA CCCTTATCCG CGCTTTCGCC ATTTGCCCTT
          201 TCAGGCTGCG GGCATAGGGG CGGAACAGGC GGCGGTCGAA TCCTGTTTCA
          251 TCCGGACAAA CGCGTTGGCA GTCGGAAAAT CCGGCCGGCC GTGTCAAATA
          301 ATGCGTTACT TTGGCCGGGT CTTGTCCTTT GTAAGCGGCG GTCTTTTTT
          351 GCGCGCCATC CGCATCTGTT TGGGCGCATG GCAAACGGCG GCTGCCGTAC
          401 AATCAAAATG TTTGGCGATT TCATGCAGAC AGGCATCCGG ATGCCGCCCG
          451 ACATATCGAG CCGGTTTTTG CCTATCCGAT TTGGCGGCAT TTAGGCCGGT
          501 AACTTGA
This corresponds to the amino acid sequence <SEQ ID 494; ORF 129.ng>:
     q129.pep
            1 MLSPPRRKTA AHQSSRLSFA CGKNAACCRD QNQYRAASSP NRGLPRFPIT
           51 PTAAAVHPYP RFRHLPFQAA GIGAEQAAVE SCFIRTNALA VGKSGRPCQI
          101 MRYFGRVLSF VSGGLFLRAI RICLGAWQTA AAVQSKCLAI SCRQASGCRP
          151 TYRAGFCLSD LAAFRPVT*
The following partial DNA sequence was identified in N. meningitidis <SEO ID 495>:
     m129.seq (partial)
            1 ..TATCTGCGCT TTCACTATTT GCCCTTTCAG GCTGCGGGCA TAGGGACGGA
                 ACAGGTAGCG GTCAAATCCT GTTTCATCCA AATAAACACG TTGGTAGTCG
           51
                 GAAAATTCGG CCGGCTGTGT CAAATAATGC GTTACTTTGG CCGGGTCTTG
          101
                 TTCTTTGTAA GTGGTGGTCT TTTTTTGCGC GTTATCCCCA TCTGTTTGAG
```

This corresponds to the amino acid sequence <SEQ ID 496; ORF 129>:

TCCGATTTGA CGGCATTTAG ACCGGTAACT TGA

m129.pep (partial)

1 ..YLRFHYLPFQ AAGIGTEQVA VKSCFIQINT LVVGKFGRLC QIMRYFGR<u>VL</u>

TGCATAGCAA ATGGTGGCTG CCGTACAATC AAAATGTTTG GCGATTTCAT GCAGATAGGC ATCCGGGTGT TGCCCAACAT ATTGAGCCGG TTTTTGCCTA

- 51 FFVSGGLFLR VIPICLSAXQ MVAAVQSKCL AISCRXASGC CPTYXAGFCL
- 101 SDLTAFRPVT *

Computer analysis of this amino acid sequence gave the following results:

BNSDOCID: <WO___9957280A2_j_>



## Homology with a predicted ORF from N. gonorrhoeae

ORF 129 shows 79.1% identity over a 110 aa overlap with a predicted ORF (ORF 129.ng) from N. gonorrhoeae:

m129/g129

m129.pep		,	YLRFH	10 YLPFQAAGIG	20 TEQVAVKSCF	30 TOTOT'
g129	RDONOYRAASSI 30 40	PNRGLPRFPITPTA 50	1 11:	:		1. 1
m129.pep g129	1 • 1 1 1 1 1 1 1	MRYFGRVLSFVSG		-1 1 -111	1111111111	1111
m129.pep g129	100 CPTYXAGFCLSD           RPTYRAGFCLSD	1:111111				
	150 160					

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 497>:
a129.seq (partial)

```
1 TATCTGCGCT TTCACTATTT GCCCTTTCAG GCTGCGGCA TAGGGACGGA
51 ACAGGTAGCG GTCAAATCCT GTTTCATCCA AATAAACACG TTGGTAGTCG
101 GAAAATCGG CCAGCTGTGT CAAATAATGC GTTACTTTGG CCGGGTCTTG
151 TTCTTTGTAA GTGGTGGTCT TTTTTTGCGC GTTATCCCCA TCTGTTTGAG
201 TGCATAGCAA ATGGTGGCTG CCGTACAATC AAAATGTTTG GCGATTTCAT
251 GCAGATAGGC ATCCTGGTGT TGCCCAACAT ATTGAGCCGG TTTTTGCCTA
301 TCCGATTTGA CGGCATTTAG ACCGGTAACT TGA
```

# This corresponds to the amino acid sequence <SEQ ID 498; ORF 129.a>: a129.pep (partial)

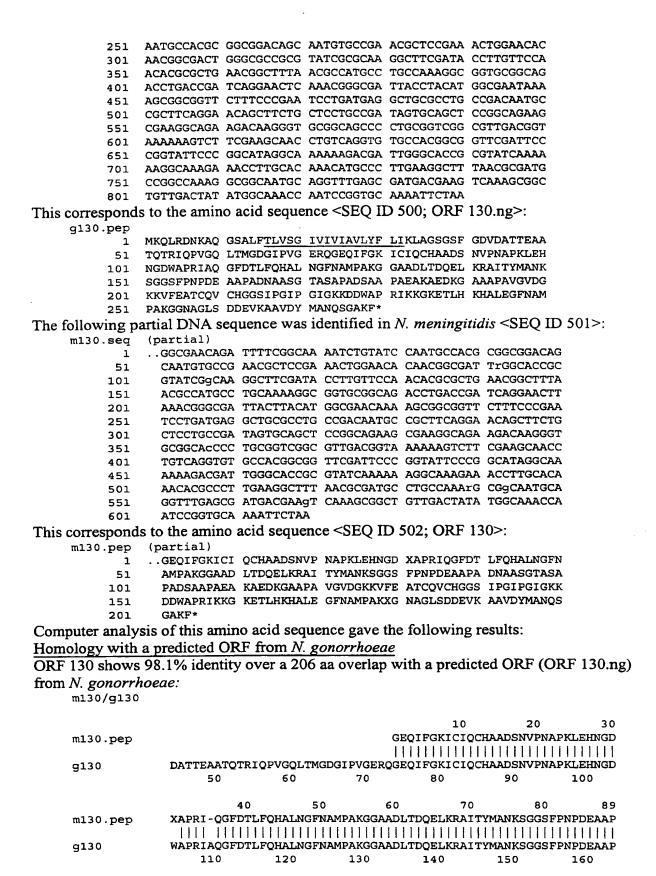
1 YLRFHYLPFQ AAGIGTEQVA VKSCFIQINT LVVGKFGQLC QIMRYFGRVL
51 FFVSGGLFLR VIPICLSA*Q MVAAVQSKCL AISCR*ASWC CPTY*AGFCL
101 SDLTAFRPVT *

## m129/a129 98.2% identity in 110 aa overlap

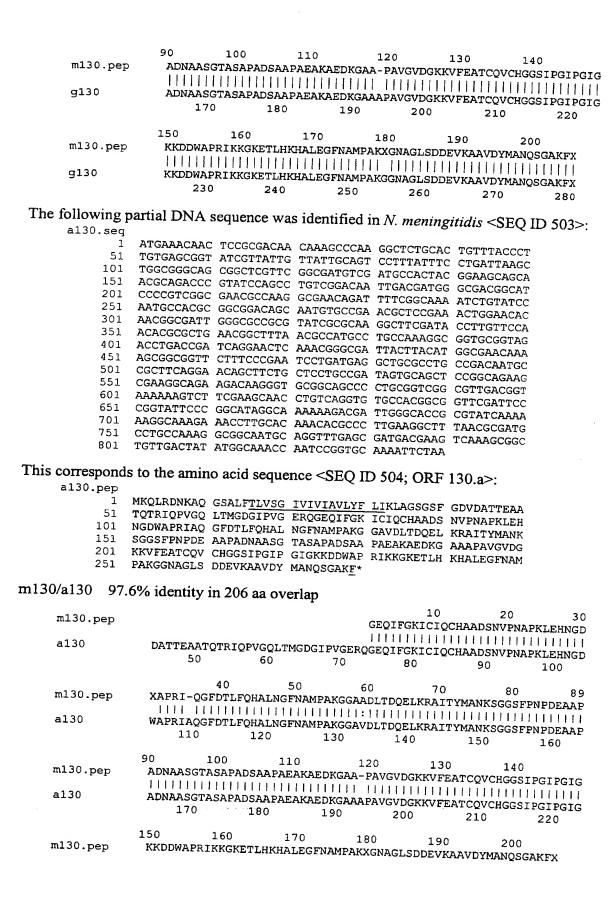
m120	10	20	30	40	50	60
m129.pep	YLRFHYLPFQAAGI	GTEQVAVKSO	FIQINTLVVO	KFGRLCOTM	VECTUL PRICE	
	11111111111	1 1 1 1 1 1 1 1 1 1	111111111	111011111		
a129	YLRFHYLPFQAAGI	GTEOVAVKSC	FTOTNTIANG	TILLITIES COLMI	VECTITE TO	11111
	10		5 TM 1 TI A A G	WE GOT COTAIN	CYFGRVLFFVS	GGLFLR
	10	20	30	40	50	60
	70	80	90	100	110	
m129.pep	VTPTCT.SAYOMUAAT	TOCKET DEDE	50	100	110	
<u>-</u>	VIPICLSAXOMVAA	\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \	RXASGCCPTY	XAGFCLSDLT	AFRPVTX	
	TIPICI SAVORIA		TILL TILL	LETTERIN	111111	
a129	VIPICLSAXOMVAA	TOCKET ATEC	D17 G77 G G G G G G G G G G G G G G G G G	,,,,,,,,,,,	111111	
	ATT TODOMNOHVAM	OSVCTUTZO	RXASWCCPTY	${ t XAGFCLSDLT}$	AFRPVTX	
	70	80	90	100	110	
			-		110	

The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 499>:

```
1 ATGAAACAAC TCCGCGACAA CAAAGCCCAA GGCTCTGCAC TGTTTACCCT
51 TGTGAGCGGT ATCGTTATTG TTATTGCAGT CCTTTATTTC CTGATTAAGC
101 TGGCGGGCAG TGGATCGTTC GGCGATGTCG ATGCCACTAC GGAAGCGGCA
151 ACGCAGACCC GCATCCAGCC TGTCGGACAA TTGACGATGG GTGACGGCAT
201 CCCCGTCGGC GAACGCCAAG GCGAACAGAT TTTCGGCAAA ATCTGTATCC
```



BNSDOCID: <WO___9957280A2_I_>



```
a130
                   KKDDWAPRIKKGKETLHKHALEGFNAMPAKGGNAGLSDDEVKAAVDYMANQSGAKFX
                                  240
                                            250 260
The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 505>:
      g132.seq
             1 ATGGAAGCCT TCAAAACCCT AATTTGGATT ATTAATATTA TTTCCGCTTT
            51 GGCCGTCATC GTGTTAGTAT TGCTCCAACA CGGCAAAGGC GCGGATGCCG
           101 GCGCGACCTT CGGATCGGGA AGCGGCAGCG CGCAAGGCGT ATTCGGCTCT
           151 GCCGGCAACG CCAACTTCCT CAGCCGCTCG ACCGCCGTTG CAGCAACATT
           201 tttcttTGca acctgcAtgg gctatggTgt atattcacac CCACACGACA
           251 AAACACGGTT TGGACTtcag caacataCGA CAGACTCAGC AagcACCCAA
301 ACCcgtAAGC AATACCGAAC CTTCTGCCCC TGTTCCTCAG CAGCAGAAAT
351 AACagtTTTT CAAATgccga caTGgtga
This corresponds to the amino acid sequence <SEQ ID 506; ORF 132.ng>:
     g132.pep
             1 MEAFKTLIWI INIISALAVI VLVLLQHGKG ADAGATFGSG SGSAQGVFGS
            51 AGNANFLSRS TAVAATFFFA TCMGYGVYSH PHDKTRFGLQ QHTTDSASTQ
           101
               TRKQYRTFCP CSSAAEITVF QMPTW*
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 507>:
     m132.seq (partial)
             1 ATGGAACCCT TCAAAACCTT AATTTGGATT GTTAATTTAA TTTCCGCTTT
            51 GGCCGTCTTC GTGTTAGTAT TGCTCCAACA CGGCAAAGGC GCGGATGCCG
           101 GCGCGACTTT CGGA...
This corresponds to the amino acid sequence <SEQ ID 508; ORF 132>:
     m132.pep
                (partial)
                MEPFKTLIWI VNLISALAVF VLVLLQHGKG ADAGATFG...
Computer analysis of this amino acid sequence gave the following results:
Homology with a predicted ORF from N. gonorrhoeae
ORF 132 shows 89.5% identity over a 38 aa overlap with a predicted ORF (ORF 132.ng)
from N. gonorrhoeae:
     m132/g132
                           10
                                      20
                                                30
                   MEPFKTLIWIVNLISALAVFVLVLLQHGKGADAGATFG
     m132.pep
                   g132
                   MEAFKTLIWIINIISALAVIVLVLLQHGKGADAGATFGSGSGSAOGVFGSAGNANFLSRS
                                      20
                                                30
                                                           40
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 509>:
     a132.seq
             1 ATGGAAGCCT TCAAAACCCT AATTTGGATT GTTAATATAA TTTCCGCTTT
           51 GGCCGTCATC GTGTTAGTAT TGCTCCAACA CGGCAAAGGC GCGGATGCCG
101 GCGCGACTTT CGGATCGGGA AGCGGCAGCG CGCAAGGCGT ATTCGGCTCT
                GCGCGACTTT CGGATCGGGA AGCGGCAGCG CGCAAGGCGT ATTCGGCTCT
           151 GCCGGCAACG CTAACTTCCT CAGCCGCTCG ACCGCCGTTG CAGCAACATT
           201 TTTCTTTGCA ACCTGCATGG GCTATGGTGT ATATTCACAC CCACACGACA
           251 AAACACGGTT TGGACTTCAG CAACGTACAA CAAACTCAGC AAGCACCCAA
           301 ACCCGTAAGC AATACCGAAC CTTCTGCCCC TGTTCCTCAG CAGCAGAAAT 351 AACAGTTTTT CAAATGCCGA CATGGTGA
This corresponds to the amino acid sequence <SEQ ID 510; ORF 132.a>:
     al32.pep
                MEAFKTLIWI VNIISALAVI VLVLLQHGKG ADAGATFGSG SGSAQGVFGS
                AGNANFLSRS TAVAATFFFA TCMGYGVYSH PHDKTRFGLQ QRTTNSASTQ
              TRKQYRTFCP CSSAAEITVF QMPTW*
m132/a132 92.1% identity in 38 aa overlap
```

201

501

```
10
                                  20
                                           30
     m132.pep
                 MEPFKTLIWIVNLISALAVFVLVLLQHGKGADAGATFG
                 MEAFKTLIWIVNIISALAVIVLVLLQHGKGADAGATFGSGSGSAQGVFGSAGNANFLSRS
     a132
                                  20
                                           30
                                                    40
                                                              50
The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 511>:
     g134.seq
              ATGTCCCAAG AAATCCTCGA CCAAGTGCGC CGCCGCCGCA CGTTTGCCAT
           1
              CATCTCCCAC CCCGATGCGG GTAAAACCAC GCTGACCGAA AAACTGCTGC
              TGTTTTCGGG CGCGATTCAA AGCGCAGGCA CGGTGAAAGG TAAGAAAACC
```

GGCAAATTCG CCACCTCCGA CTGGATGGAC ATCGAGAAGC AGCGCGGCAT

TTCCGTGGCA TCAAGCGTGA TGCAGTTCGA CTACAAAGAC CACACCGTCA ACCTCTTGGA CACGCCGGGA CACCAAGACT TCTCCGAAGA CACCTACCGC GTTTTAACCG CAGTGGACAG CGCCTTGATG GTCATCGACG CGGCAAAAGG CGTGGAAGCG CAAACCATCA AACTCTTGAA CGTCTGCCGC CTGCGCGATA CGCCGATTGT TACCTTCATG AACAAATACG ACCGCGAAGT GCGCGATTCT TTGGAACTCT TGGACGAAGT GGAAGACATC CTGCAAATCC GCTGCGCGCC 501 CGTTACCTGG CCGATCGGTA TGGGCAAAAA CTTCAAGGGC GTGTACCACA 551 TCCTGAACGA CGAAATCTAT CTCTTTGAAG CGGGCGGCGA ACGCCTGCCG 601 CACGAGTTCG ACATCATCAA AGGCATAAAC AATCCCGAAT TGGAACAACG CTTTCCGTTG GAAATCCAGC AGTTGCGCGA CGAAATCGAA TTGGTGCAGG 651 CGGCTTCCAA CGAATTTAAT CTCGacgaAT TTCTCGccgG CGAACTCACG 701 CCAGTGTTCT TCGGCTCTGC GATTAACAAC TTCGGCATTC AGGAAATCCT 751 801 CAATTCATTG ATTGACTGGG CACCCGCACC GAAACCGCGC GACGCGACCA 851 TGCGCATGGT CGGGCCGGAC GAGCCGAAAT TTTCCGGATT TATCTTTAAA 901 ATCCAAGCCA ATATGGACCC GAAACACCGC GACCGTATCG CCTTCTTGCG 951 CGTCTGCTCC GGTAAATTCG AGCGCGGCAT GAAGATGAAA CACCTGCGTA 1001 TCAACCGCGA AATCGCCGCC TCCAGCGTAG TAACCTTCAT GTCGCACGAC 1051 CGCGAACTGG CGGAAGAAGC CTACGCCGGC GACATCATCG GCATCCCGAA 1101 CCACGGCAAC ATCCAAATCG GCGACAGCTT CTCCGAAGGC GAACAACTGG 1151 CGTTTACCGG CATCCCATTC TTCGCGCCCG AACTGTTCCG CAGCGTCCGC 1201 ATCAAAAACC CGCTGAAAAT CAAACAACTG CAAAAAGGTT TGCAACAACT CGGCGAAGAA GGTGCGGTTC AAGTATTCAA ACCGATGAGC GGCGCGGATT

TGATTTTGGG TGCGGTCGGC GTGTTGCAGT TTGAAGTCGT AACCTCACGC CTCGCCAACG AATACGGCGT GGAAGCCGTG TTCGACAGCG CATCCATCTG 1401 GTCGGCGCGC TGGGTATCGT GCGACGACAA GAAAAACTG GCGGAATTTG 1451 AAAAAGCCAA CGCAGGCAAC CTCGCCATCG ACGCAGGCGG CAACCTCGCC 1501 TACCTCGCCC CCAACCGCGT GAATTTGGGG TTGACGCAAG AACGCTGGCC

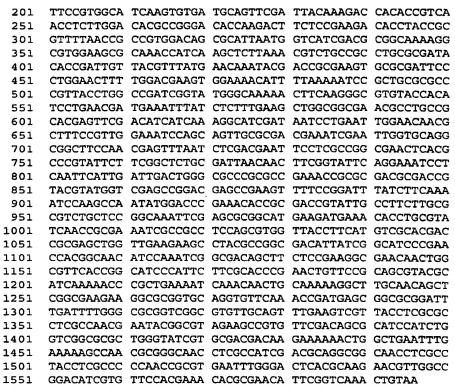
1551 GGACATCGTG TTCCACGAAA CGCGCGAACA TTCGGTCAAA CTCTAA This corresponds to the amino acid sequence <SEQ ID 512; ORF 134.ng>:

YLAPNRVNLG LTQERWPDIV FHETREHSVK L*

g134.pep MSQEILDQVR RRRTFAIISH PDAGKTTLTE KLLLFSGAIQ SAGTVKGKKT GKFATSDWMD IEKQRGISVA SSVMQFDYKD HTVNLLDTPG HQDFSEDTYR 51 VLTAVDSALM VIDAAKGVEA QTIKLLNVCR LRDTPIVTFM NKYDREVRDS 151 LELLDEVEDI LQIRCAPVTW PIGMGKNFKG VYHILNDEIY LFEAGGERLP 201 HEFDIIKGIN NPELEQRFPL EIQQLRDEIE LVQAASNEFN LDEFLAGELT 251 PVFFGSAINN FGIQEILNSL IDWAPAPKPR DATMRMVGPD EPKFSGFIFK IQANMDPKHR DRIAFLRVCS GKFERGMKMK HLRINREIAA SSVVTFMSHD 351 RELAEEAYAG DIIGIPNHGN IQIGDSFSEG EQLAFTGIPF FAPELFRSVR 401 IKNPLKIKQL QKGLQQLGEE GAVQVFKPMS GADLILGAVG VLQFEVVTSR 451 LANEYGVEAV FDSASIWSAR WVSCDDKKKL AEFEKANAGN LAIDAGGNLA

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 513>: m134.seg

> ATGTCCCAAG AAATCCTCGA CCAAGTGCGC CGCCGCCGCA CGTTTGCCAT 51 CATCTCCCAC CCTGACGCAG GTAAAACCAC GTTGACTGAA AAACTCTTGC 101 TGTTTTCGGG CGCGATTCAG AGCGCGGGTA CGGTAAAAGG CAAGAAAACC GGCAAATTCG CCACTTCCGA CTGGATGGAA ATCGAGAAGC AGCGCGGCAT



This corresponds to the amino acid sequence <SEQ ID 514; ORF 134>:

```
m134.pep

1 MSQEILDQVR RRRTFAIISH PDAGKTTLTE KLLLFSGAIQ SAGTVKGKKT
51 GKFATSDWME IEKQRGISVA SSVMQFDYKD HTVNLLDTPG HQDFSEDTYR
101 VLTAVDSALM VIDAAKGVEA QTIKLLNVCR LRDTPIVTFM NKYDREVRDS
151 LELLDEVENI LKIRCAPVTW PIGMGKNFKG VYHILNDEIY LFEAGGERLP
201 HEFDIIKGID NPELEQRFPL EIQQLRDEIE LVQAASNEFN LDEFLAGELT
251 PVFFGSAINN FGIQEILNSL IDWAPAPKPR DATVRMVEPD EPKFSGFIFK
301 IQANMDPKHR DRIAFLRVCS GKFERGMKMK HLRINREIAA SSVVTFMSHD
351 RELVEEAYAG DIIGIPNHGN IQIGDSFSEG EQLAFTGIPF FAPELFRSVR
401 IKNPLKIKQL QKGLQQLGEE GAVQVFKPMS GADLILGAVG VLQFEVVTSR
451 LANEYGVEAV FDSASIWSAR WVSCDDKKKL AEFEKANAGN LAIDAGGNLA
```

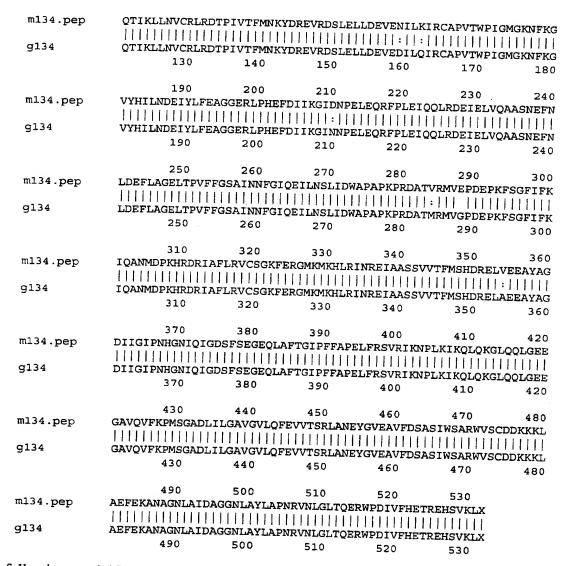
Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae

ORF 134 shows 98.7% identity over a 531 aa overlap with a predicted ORF (ORF 134.ng) from N. gonorrhoeae:

m134/g134

	10	20	30	40	50	60
m134.pep	MSQEILDQVRRRRT	FAIISHPDAG	KTTLTEKLLL	FSGAIQSAGT	VKGKKTGKFA	TSDWME
			1111111111		1111111111	11111:
g134	MSQEILDQVRRRRT	FAIISHPDAG	KTTLTEKLLL	FSGAIQSAGT	VKGKKTGKFA	TSDWMD
	10	20	30	40	50	60
	70	80	90	100	110	120
m134.pep	IEKQRGISVASSVM	QFDYKDHTVN	LLDTPGHQDF	SEDTYRVLTA	VDSALMVIDA	AKGVEA
			1111111111	11111111111	1111111111	
g134	IEKQRGISVASSVM	QFDYKDHTVN	LLDTPGHQDF	SEDTYRVLTA	VDSALMVIDA	AKGVEA
	70	80	90	100	110	120
	130	140	150	160	170	180



# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 515>:

1	ATGTCCCAAG	AAATCCTCGA	CCAAGTGCGC	CGCCGCCGCA	CGTTTGCCAT
51	CATCTCCCAC	CCTGACGCAG	GTAAAACCAC	GTTGACTGAA	
101	TGTTTTCAGG	TGCGATTCAA			
151	GGCAAATTCG	CCACCTCCGA			
201	TTCCGTGGCA	TCAAGCGTGA	TGCAGTTCGA	411 0 011 01 W 70 C	
251	ACCTTTTGGA		CACCAAGACT	4-11177710110	
301	GTTTTGACCG		TGCCTTGATG	TCTCCGAAGA	
351	CGTGGAAGCG		AACTCTTGAA	GTCATCGACG	
401	CGCCGATTGT		AACAAATACG	CGTCTGCCGC	or o o o o o i i i i i
451	CTGGAATTGC	TGGACGAAGT	GGAAAACATC	ACCGCGAAGT	GCGCGATTCC
501	CGTAACCTGG		TGGGCAAAAA	CTGCAAATCC	GCTGCGCGCC
551	TCCTGAACGA		CTCTTTGAAG	CTTCAAAGGC	GTGTACCACA
601	CACGAGTTCG		AGGCATCGAT	CGGGCGGCGA	
651	CTTTCCGTTA	*******	AGTTGCGCGA	AATCCCGAAT	TGGAACAACG
701	CGGCTTCCAA	0.111101100		CGAAATCGAA	TTGGTGCAGG
751		TCGGCTCTGC	CTCGACGAAT	TCCTCGCCGG	CGAACTCACG
801	CAATTCATTG	ATTGAATGGG	GATTAACAAC	TTCGGTATTC	AGGAAATCCT
851	TGCGTATGGT		CGCCCGCGCC	GAAACCACGC	GATGCGACCG
901	ATCCAAGCCA	CGAGCCGGAC	~~~~~	TTTCCGGATT	TATCTTCAAA
	COLINGCOM	ATATGGACCC	GAAACACCGC	GACCGTATTG	CCTTCTTGCG

951	CGTCTGCTCC	GGCAAATTCG	AGCGCGGCAT	GAAAATGAAA	CACCTGCGTA	
1001	TCAACCGCGA	AATCGCCGCC	TCCAGCGTGG	TAACCTTCAT	GTCCCACGAC	
1051			CTACGCCGGC			
1101			GCGACAGCTT			
1151	CGTTTACCGG	CATCCCATTC	TTCGCGCCCG	AACTGTTCCG	CAGCGTTCGC	
1201	ATCAAAAACC	CGCTGAAAAT	CAAGCAACTG	CAAAAAGGTT	TGCAACAGCT	
1251			AGGTGTTCAA			
1301			GTGTTGCAGT			
1351	CTTGCCAACG	AATACGGCGT	GGAAGCCGTG	TTCGACAACG	CATCCATCTG	
1401	GTCGGCGCGC	TGGGTATCGT	GCGACGACAA	GAAAAAACTG	GCGGAATTTG	
1451			CTCGCCATCG			
1501			GAATCTGGGA			
1551	GGACATCGTG	TTCCACGAAA	CGCGCGAGCA	TTCGGTCAAA	CTTTAA	
This correspond	s to the amin	o acid seque	nce <seo ii<="" td=""><td>7 516 ORF</td><td>134 2&gt;</td><td></td></seo>	7 516 ORF	134 2>	
•	is to the airmin	o acia seque	wee ord u	5 510, Old	134.00.	
a134.pep						
1	MSQEILDQVR	RRRTFAIISH	PDAGKTTLTE	KLLLFSGAIO	SAGTVKGKKT	
51			SSVMQFDYKD			
101			QTIKLLNVCR			
151			PIGMGKNFKG			
201	HEFDIIKGID	NPELEORFPL	EIQQLRDEIE	LVOAASNEFN	LDEFLAGELT	
251			IEWAPAPKPR			
301			GKFERGMKMK			
351			IQIGDSFSEG			
401	IKNPLKIKOL	OKGLOOLGEE	GAVQVFKPMS	GADLILGAVG	VLOFEVVTSR	
451			WVSCDDKKKL			
					DAIDAGGNDA	
501	YLAPNKVNLG	LIGERMADIA	FHETREHSVK	μ <b>.</b>		
m134/a134 98	.9% identity	in 531 aa ov	erlap			
	•		20 2	10	E 0	<b>CO</b> :
					50	60
m134.pep	MSQEILDÇ	VRRRRTFAII:	SHPDAGKTTLTI	EKLLLFSGAIQS	SAGTVKGKKTGK	FATSDWME
			11111111			111111:
a134				HILLESGATOS		
a134		VRRRRTFAII:	SHPDAGKTTLT	EKLLLFSGAIQS	SAGTVKGKKTGK	FATSDWMD
a134		VRRRRTFAII:		EKLLLFSGAIQS		
a134		OVRRRRTFAIIS 10	SHPDAGKTTLTE 20 30	EKLLLFSGAIQS 0 40	SAGTVKGKKTGK 50	FATSDWMD 60
	MSQEILDQ	OVRRRRTFAIIS 10 2 70 8	SHPDAGKTTLTE 20 30 80 90	EKLLLFSGÄIQS 0 40 0 100	SAGTVKGKKTGK 50 110	FATSDWMD 60
	MSQEILDQ	OVRRRRTFAIIS 10 2 70 8	SHPDAGKTTLTE 20 30 80 90	EKLLLFSGÄIQS 0 40 0 100	SAGTVKGKKTGK 50	FATSDWMD 60
a134 m134.pep	MSQEILDÇ IEKQRGIS	OVRRRRTFAIIS 10 2 70 8 SVASSVMQFDY	SHPDAGKTTLTE 20 30 80 90 KDHTVNLLDTPO	EKLLLFSGAIQS ) 40 ) 100 GHQDFSEDTYRV	SAGTVKGKKTGK 50 110 VLTAVDSALMVI	FATSDWMD 60 120 DAAKGVEA
m134.pep	MSQEILDQ IEKQRGIS 	70 8 VASSVMQFDYF	SHPDAGKTTLTE 20 30 80 90 KDHTVNLLDTPO	EKLLLFSGAIQS 0 40 0 100 SHQDFSEDTYRV	SAGTVKGKKTGK 50 110 VLTAVDSALMVI	FATSDWMD 60 120 DAAKGVEA
	MSQEILDQ IEKQRGIS 	OVRRRRTFAIIS 10 8 70 8 VVASSVMQFDYF	SHPDAGKTTLTI 20 30  80 90  KDHTVNLLDTPO	EKLLLFSGAIQS  100  HQDFSEDTYR  HIIIIIIIII	SAGTVKGKKTGK 50 110 VLTAVDSALMVI	FATSDWMD 60 120 DAAKGVEA 
m134.pep	MSQEILDQ IEKQRGIS 	OVRRRRTFAIIS 10 8 70 8 VVASSVMQFDYF	SHPDAGKTTLTE 20 30 80 90 KDHTVNLLDTPO	EKLLLFSGAIQS  100  HQDFSEDTYR  HIIIIIIIII	SAGTVKGKKTGK 50 110 VLTAVDSALMVI	FATSDWMD 60 120 DAAKGVEA
m134.pep	MSQEILDQ IEKQRGIS         IEKQRGIS	OVRRRRTFAIIS 10 8 70 8 SVASSVMQFDYR !!!!!!!!	SHPDAGKTTLTI 20 30  80 90  KDHTVNLLDTPO	EKLLLFSGAIQS  100  HQDFSEDTYR  HIIIIIIIII	SAGTVKGKKTGK 50 110 VLTAVDSALMVI	FATSDWMD 60 120 DAAKGVEA 
m134.pep	MSQEILDQ IEKQRGIS         IEKQRGIS	OVRRRRTFAIIS 10 8 70 8 SVASSVMQFDYR           SVASSVMQFDYR	SHPDAGKTTLTI 20 30  80 90  KDHTVNLLDTPO	EKLLLFSGAIQS  100  SHQDFSEDTYRV  SHQDFSEDTYRV  GHQDFSEDTYRV  HILLIHIHI  HQDFSEDTYRV  100	SAGTVKGKKTGK 50 110 VLTAVDSALMVI	FATSDWMD 60 120 DAAKGVEA 
m134.pep a134	MSQEILDQ IEKQRGIS         IEKQRGIS	TO ESTABLE TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STANDARD TO STAN	SHPDAGKTTLTI 20 30  80 90  KDHTVNLLDTPO	EKLLLFSGAIQS  100  SHQDFSEDTYRV  SHQDFSEDTYRV  CHQDFSEDTYRV  CHQDFSEDTYRV  100  100	SAGTVKGKKTGK 50 110 YLTAVDSALMVI                     YLTAVDSALMVI 110	FATSDWMD 60 120 DAAKGVEA         DAAKGVEA 120 180
m134.pep	MSQEILDO IEKQRGIS IIIIIIIII IEKQRGIS 1	OVRRRRTFAIIS 10 8 70 8 SVASSVMQFDYF           SVASSVMQFDYF 70 8	### SHPDAGKTTLTI    1	EKLLLFSGAIQS  100  SHQDFSEDTYRV  SHQDFSEDTYRV  1111111111111111  SHQDFSEDTYRV  100  160  SLELLDEVENII	SAGTVKGKKTGK 50 110 /LTAVDSALMVI                       /LTAVDSALMVI 110 170 LKIRCAPVTWPI	FATSDWMD 60 120 DAAKGVEA         DAAKGVEA 120 180 GMGKNFKG
m134.pep a134 m134.pep	MSQEILDO IEKQRGIS IIIIIIII IEKQRGIS 1 QTIKLLNV	OVRRRRTFAIIS 10 2 70 8 SVASSVMQFDYB                     SVASSVMQFDYB 70 8 .30 14 VCRLRDTPIVTB	SHPDAGKTTLTE 20 30  80 90  KDHTVNLLDTPO                      KDHTVNLLDTPO 80 90  40 150  FMNKYDREVRDS	EKLLLFSGAIQS  40  100  SHQDFSEDTYRV  SHQDFSEDTYRV  100  160  SLELLDEVENII	SAGTVKGKKTGK 50  110  */LTAVDSALMVI                      */LTAVDSALMVI 110  170  .KIRCAPVTWPI	FATSDWMD 60 120 DAAKGVEA         DAAKGVEA 120 180 GMGKNFKG
m134.pep a134	MSQEILDO IEKQRGIS IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	70 8  VASSVMQFDYF  (1) (1) (1) (1) (1) (2) (3) (2) (4) (4) (4) (4) (4) (4) (4) (4) (4) (4	######################################	EKLLLFSGAIQS  100  GHQDFSEDTYRV  GHQDFSEDTYRV  1111111111111111  GHQDFSEDTYRV  100  160  SLELLDEVENII	SAGTVKGKKTGK 50 110 /LTAVDSALMVI                       /LTAVDSALMVI 110 170 LKIRCAPVTWPI	FATSDWMD 60 120 DAAKGVEA         DAAKGVEA 120 180 GMGKNFKG
m134.pep a134 m134.pep	MSQEILDO IEKQRGIS IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	70 8  VASSVMQFDYF  (1) (1) (1) (1) (1) (2) (3) (2) (4) (4) (4) (4) (4) (4) (4) (4) (4) (4	SHPDAGKTTLTE 20 30  80 90  KDHTVNLLDTPO                      KDHTVNLLDTPO 80 90  40 150  FMNKYDREVRDS	EKLLLFSGAIQS  100  GHQDFSEDTYRV  GHQDFSEDTYRV  1111111111111111  GHQDFSEDTYRV  100  160  SLELLDEVENII	SAGTVKGKKTGK 50  110  */LTAVDSALMVI                      */LTAVDSALMVI 110  170  .KIRCAPVTWPI	FATSDWMD 60 120 DAAKGVEA         DAAKGVEA 120 180 GMGKNFKG
m134.pep a134 m134.pep	MSQEILDO IEKQRGIS IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	70 8  VASSVMQFDYF  (1) (1) (1) (1) (1) (2) (3) (2) (4) (4) (4) (4) (4) (4) (4) (4) (4) (4	######################################	EKLLLFSGAIQS  100  GHQDFSEDTYRV  GHQDFSEDTYRV  1111111111111111  GHQDFSEDTYRV  100  160  SLELLDEVENII	SAGTVKGKKTGK 50  110  /LTAVDSALMVI          /LTAVDSALMVI 110  170  LKIRCAPVTWPI :	FATSDWMD 60  120 DAAKGVEA         DAAKGVEA 120  180 GMGKNFKG
m134.pep a134 m134.pep	MSQEILDO  IEKQRGIS  IIIIIIII IEKQRGIS  QTIKLLNU IIIIIIIII	70 8 5VASSVMQFDYE 11111111115 5VASSVMQFDYE 70 8 .30 14 7CRLRDTPIVTE 111111117 7CRLRNTPIVTE .30 14	######################################	EKLLLFSGAIQS  40  100  SHQDFSEDTYR  SHQDFSEDTYR  100  160  SLELLDEVENII  SLELLDEVENII  1111111111111111111111111111111111	SAGTVKGKKTGK 50  110  */LTAVDSALMVI            */LTAVDSALMVI 110  170  .KIRCAPVTWPI :         .QIRCAPVTWPI	FATSDWMD 60  120 DAAKGVEA         DAAKGVEA 120  180 GMGKNFKG
m134.pep a134 m134.pep a134	MSQEILDO  IEKQRGIS            IEKQRGIS  1  QTIKLLNV            QTIKLLNV	70 8 8VASSVMQFDYI 10 11 11 11 11 11 11 11 11 11 11 11 11 1	SHPDAGKTTLTE 20 30  80 90  KDHTVNLLDTPO 111111111111111111111111111111111111	EKLLLFSGAIQS  40  100  GHQDFSEDTYRV  SHQDFSEDTYRV  100  160  SLELLDEVENII	SAGTVKGKKTGK 50  110  /LTAVDSALMVI                     /LTAVDSALMVI 110  170  LKIRCAPVTWPI :                   LQIRCAPVTWPI 170  230	TATSDWMD 60  120 DAAKGVEA IIIIIIII DAAKGVEA 120  180 GMGKNFKG IIIIIIII GMGKNFKG 180 240
m134.pep a134 m134.pep	MSQEILDO  IEKQRGIS  IIIIIIIII  QTIKLLNU  QTIKLLNU  1  VYHILNDE	TO SEVASSVMQFDYN SVASSVMQFDYN SVASSVMQFDYN TO SEVASSVMQFDYN SHPDAGKTTLTE 20 30  80 90  KDHTVNLLDTPO 111111111111111111111111111111111111	EKLLLFSGAIQS  100  GHQDFSEDTYRV  SHQDFSEDTYRV  11111111111111111111111111111111111	SAGTVKGKKTGK 50  110  VLTAVDSALMVI 110  170  LKIRCAPVTWPI 170  230  CIQQLRDEIELV	TATSDWMD 60  120 DAAKGVEA         DAAKGVEA 120  180 GMGKNFKG         GMGKNFKG 180 240 QAASNEFN	
m134.pep a134 m134.pep a134	MSQEILDO  IEKQRGIS            IEKQRGIS  1  QTIKLLNU            QTIKLLNU  1  VYHILNDE	OVRRRRTFAIIS 10 70 8VASSVMQFDYI           8VASSVMQFDYI 70 8 .30 14 VCRLRDTPIVTI 7CRLRNTPIVTI 7CRLRNTPIVTI 30 14 .90 20 EIYLFEAGGERI	SHPDAGKTTLTE 20 30  80 90  KDHTVNLLDTPO 111111111111111111111111111111111111	EKLLLFSGAIQS  100  GHQDFSEDTYRV  SHQDFSEDTYRV  1111111111111111  SHELLDEVENII  SLELLDEVENII  SLELLDEVENII  11111111111111  220  DNPELEQRFPLE	SAGTVKGKKTGK 50  110  VLTAVDSALMVI 110  170  LKIRCAPVTWPI 170  230  CIQQLRDEIELV	TATSDWMD 60  120 DAAKGVEA         DAAKGVEA 120  180 GMGKNFKG          GMGKNFKG 180  240 QAASNEFN
m134.pep a134 m134.pep a134	MSQEILDO  IEKQRGIS  IIIIIIIII  QTIKLLNU  IIIIIIIII  T  VYHILNDE  IIIIIIIII  VYHILNDE	OVRRRRTFAIIS 10 70 8VASSVMQFDYI            8VASSVMQFDYI 70 8 .30 14 .7CRLRDTPIVTI 7CRLRNTPIVTI 7CRLRNTPIVTI 30 14 .90 20 EIYLFEAGGERI 1	SHPDAGKTTLTE 20 30  80 90  KDHTVNLLDTPO 111111111111111111111111111111111111	EKLLLFSGAIQS  100  GHQDFSEDTYRV  SHQDFSEDTYRV  1111111111111111  SHELLDEVENII  SLELLDEVENII  SLELLDEVENII  11111111111111  220  DNPELEQRFPLE	SAGTVKGKKTGK 50  110  VLTAVDSALMVI 110  170  LKIRCAPVTWPI 170  230  CIQQLRDEIELV	TATSDWMD 60  120 DAAKGVEA         DAAKGVEA 120  180 GMGKNFKG          GMGKNFKG 180  240 QAASNEFN
m134.pep a134 m134.pep a134	MSQEILDO  IEKQRGIS  IIIIIIIII  QTIKLLNU  IIIIIIIII  T  VYHILNDE  IIIIIIIII  VYHILNDE	OVRRRRTFAIIS 10 70 8VASSVMQFDYI           8VASSVMQFDYI 70 8 20 11 12 130 14 14 15 16 17 17 18 19 19 19 19 19 19 19 19 19 19 19 19 19	SHPDAGKTTLTE 20 30  80 90  KDHTVNLLDTPO 111111111111111111111111111111111111	EKLLLFSGAIQS  40  100  GHQDFSEDTYRV  SHQDFSEDTYRV  100  160  SLELLDEVENIN  SLELLDEVENIN  1111111111111111111111111111111111	SAGTVKGKKTGK 50  110  VLTAVDSALMVI 110  170  LKIRCAPVTWPI 170  230  CIQQLRDEIELV	TATSDWMD 60  120 DAAKGVEA         DAAKGVEA 120  180 GMGKNFKG          GMGKNFKG 180  240 QAASNEFN
m134.pep a134 m134.pep a134	MSQEILDO  IEKQRGIS  IIIIIIIII  QTIKLLNU  IIIIIIIII  T  VYHILNDE  IIIIIIIII  VYHILNDE	OVRRRRTFAIIS 10 70 8VASSVMQFDYI           8VASSVMQFDYI 70 8 20 11 12 130 14 14 15 16 17 17 18 19 19 19 19 19 19 19 19 19 19 19 19 19	SHPDAGKTTLTE 20 30  80 90  KDHTVNLLDTPO 111111111111111111111111111111111111	EKLLLFSGAIQS  40  100  GHQDFSEDTYRV  SHQDFSEDTYRV  100  160  SLELLDEVENII	SAGTVKGKKTGK 50  110  /LTAVDSALMVI                     /LTAVDSALMVI 110  170  LKIRCAPVTWPI :                 LQIRCAPVTWPI 170  230  CIQQLRDEIELV                   CIQQLRDEIELV	TATSDWMD 60  120 DAAKGVEA IIIIIIII DAAKGVEA 120  180 GMGKNFKG IIIIIIII GMGKNFKG 180  240 QAASNEFN IIIIIIIII
m134.pep a134 m134.pep a134	MSQEILDO  IEKQRGIS            IEKQRGIS  1  QTIKLLNV            QTIKLLNV  1  VYHILNDE	70 8  VASSVMQFDYE  1111111111  VASSVMQFDYE  70 8  20 14  VCRLRDTPIVTE  1111111111  VCRLRNTPIVTE  20 20  21 14  21 14 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14 14  21 14 14  21 14 14  21 14 14  21 14 14  21 14 14  21 14 14  21 14 14  21 14 14  21 14 14  21 14 14  21 14 14  21 14 14  21 14 14  21 14 14  21 14 14  21 14 14  21 14 14  21 14 14  21 14 14  21 14 14  2	SHPDAGKTTLTE 20 30  80 90  KDHTVNLLDTPO                        KDHTVNLLDTPO 80 90  40 150  FMNKYDREVRDS                      FMNKYDREVRDS 40 150  00 210  LPHEFDIIKGII                    LPHEFDIIKGII 00 210	EKLLLFSGAIQS  100  GHQDFSEDTYRV	SAGTVKGKKTGK 50  110  VLTAVDSALMVI 110  170  KIRCAPVTWPI 170  230  IQQLRDEIELV 110	120 DAAKGVEA                 DAAKGVEA                 DAAKGVEA   120  180 GMGKNFKG                 GMGKNFKG 180  240 QAASNEFN                 QAASNEFN 240
m134.pep a134 m134.pep a134 m134.pep a134	MSQEILDO  IEKQRGIS           IEKQRGIS  1 QTIKLLNV         QTIKLLNV 1 VYHILNDE          VYHILNDE	70 8  VASSVMQFDYI  10 10  70 8  VASSVMQFDYI  70 8  20  20  20  20  21  21  21  21  21  21	SHPDAGKTTLTE 20 30 80 90 KDHTVNLLDTP0                         KDHTVNLLDTP0 80 90 40 150 FMNKYDREVRDS                         FMNKYDREVRDS 40 150 00 210 LPHEFDIIKGII LPHEFDIIKGII 00 210	EKLLLFSGAIQS  100  GHQDFSEDTYRV	SAGTVKGKKTGK 50  110  VLTAVDSALMVI 1110  170  LKIRCAPVTWPI 170  230  CIQQLRDEIELV 11QULRDEIELV 230  230  290	120 DAAKGVEA                 DAAKGVEA                 DAAKGVEA   120  180 GMGKNFKG                 GMGKNFKG   180  240  QAASNEFN                 QAASNEFN 240  300
m134.pep a134 m134.pep a134	MSQEILDO  IEKQRGIS           IEKQRGIS  1 QTIKLLNV          QTIKLLNV 1 VYHILNDE          VYHILNDE	TO SEVASSVMQFDYN SVASSVMQFDYN SVASSVMQFDYN 70 SEVASSVMQFDYN	SHPDAGKTTLTE 20 30 80 90 KDHTVNLLDTP0                         KDHTVNLLDTP0 80 90 40 150 FMNKYDREVRDS                         FMNKYDREVRDS                       FMNKYDREVRDS                         FMNKYDREVRDS	EKLLLFSGAIQS  100  GHQDFSEDTYRV	SAGTVKGKKTGK 50  110  VLTAVDSALMVI 1110  170  LITAVDSALMVI 110  270  LKIRCAPVTWPI 170  230  CIQQLRDEIELV 111111111  CIQQLRDEIELV 230  290  DATVRMVEPDEP	TATSDWMD 60  120 DAAKGVEA          DAAKGVEA 120  180 GMGKNFKG          GMGKNFKG 180  240 QAASNEFN          QAASNEFN 240 300 PKFSGFIFK
m134.pep a134 m134.pep a134 m134.pep a134	MSQEILDO  IEKQRGIS  IIIIIIII IEKQRGIS  1 QTIKLLNV IIIIIIII QTIKLLNV IIIIIIIII VYHILNDE IIIIIIIII VYHILNDE IIIIIIIII LDEFLAGE	OVRRRRTFAIIS 10 2 70 8 SVASSVMQFDYF                     SVASSVMQFDYF 70 8 .30 14 VCRLRDTPIVTI                 VCRLRNTPIVTI .30 14 .50 20 .11 .11 .11 .11 .11 .11 .11 .22 .33 .30 20 .30 .30 .30 .30 .30 .30 .30 .30 .30 .3	SHPDAGKTTLTE 20 30 80 90 KDHTVNLLDTPC                       KDHTVNLLDTPC 80 90 40 150 FMNKYDREVRDS                     FMNKYDREVRDS                     FMNKYDREVRDS                       FMNKYDREVRDS                       FMNKYDREVRDS                         FMNKYDREVRDS                         FMNKYDREVRDS	EKLLLFSGAIQS  40  100  SHQDFSEDTYRY  SHQDFSEDTYRY  100  160  SLELLDEVENII  SLELLDEVENII  1111111111111111111111111111111111	SAGTVKGKKTGK 50  110  /LTAVDSALMVI           /LTAVDSALMVI 110  170  .KIRCAPVTWPI :          .QIRCAPVTWPI 170  230  CIQQLRDEIELV           CIQQLRDEIELV           CIQQLRDEIELV           CIQQLRDEIELV	120 DAAKGVEA         DAAKGVEA          DAAKGVEA          DAAKGVEA   20  180 GMGKNFKG          GMGKNFKG 180  240 QAASNEFN          QAASNEFN 240  300 PKFSGFIFK
m134.pep a134 m134.pep a134 m134.pep a134	MSQEILDO  IEKQRGIS  IIIIIIII IEKQRGIS  1 QTIKLLNV IIIIIIII QTIKLLNV IIIIIIIII VYHILNDE IIIIIIIII VYHILNDE IIIIIIIII LDEFLAGE	OVRRRRTFAIIS 10 2 70 8 SVASSVMQFDYF                     SVASSVMQFDYF 70 8 .30 14 VCRLRDTPIVTI                 VCRLRNTPIVTI .30 14 .50 20 .11 .11 .11 .11 .11 .11 .11 .22 .33 .30 20 .30 .30 .30 .30 .30 .30 .30 .30 .30 .3	SHPDAGKTTLTE 20 30 80 90 KDHTVNLLDTPC                       KDHTVNLLDTPC 80 90 40 150 FMNKYDREVRDS                     FMNKYDREVRDS                     FMNKYDREVRDS                       FMNKYDREVRDS                       FMNKYDREVRDS                         FMNKYDREVRDS                         FMNKYDREVRDS	EKLLLFSGAIQS  40  100  SHQDFSEDTYRY  SHQDFSEDTYRY  100  160  SLELLDEVENII  SLELLDEVENII  1111111111111111111111111111111111	SAGTVKGKKTGK 50  110  /LTAVDSALMVI           /LTAVDSALMVI 110  170  .KIRCAPVTWPI :          .QIRCAPVTWPI 170  230  CIQQLRDEIELV           CIQQLRDEIELV           CIQQLRDEIELV           CIQQLRDEIELV	120 DAAKGVEA         DAAKGVEA          DAAKGVEA          DAAKGVEA   20  180 GMGKNFKG          GMGKNFKG 180  240 QAASNEFN          QAASNEFN 240  300 PKFSGFIFK
m134.pep a134 m134.pep a134 m134.pep a134	MSQEILDO  IEKQRGIS           IEKQRGIS  1 QTIKLLNV          QTIKLLNV 1 VYHILNDE          VYHILNDE    LDEFLAGE	TO SEVASSVMQFDYR SVASSVMQFDYR SVASSVMQFDYR TO SEVASSVMQFDYR  SEVASSVMQFT TO SEVASSVMQFT TO SEVASSVMQFT TO SEVASSVMQFT TO SEVASSVMQFT TO SEVASSVMQFT TO SEVASSVMQFT TO SEVASSVMQFT TO SEVASSVMQFT TO SEVASSVMQFT TO SEVASSVMQFT TO SEVASSVMQFT TO SEVASSVMQFT TO SEVASSVMQFT TO SEVASSVMQFT TO SEVASSVMQFT TO SEVASSVMQFT	SHPDAGKTTLTE 20 30 80 90 KDHTVNLLDTP0                         KDHTVNLLDTP0 80 90 40 150 FMNKYDREVRD                       FMNKYDREVRD 40 150 00 210 LPHEFDIIKGII                     LPHEFDIIKGII 00 210 NNFGIQEILNSI	EKLLLFSGAIQS  100  GHQDFSEDTYRV	SAGTVKGKKTGK 50  110  /LTAVDSALMVI	120 120 120 120 120 120 120 120 120 180 120 180 180 180 180 180 180 180 180 180 18
m134.pep a134 m134.pep a134 m134.pep a134	MSQEILDO  IEKQRGIS           IEKQRGIS  1 QTIKLLNV          QTIKLLNV 1 VYHILNDE          VYHILNDE    LDEFLAGE	70 8  VASSVMQFDYE  111111111  VASSVMQFDYE  70 8  20ASSVMQFDYE  70 8  20ASSVMQFDYE  70 8  20ASSVMQFDYE  70 8  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASSVMQFDYE  20ASS	SHPDAGKTTLTE 20 30 80 90 KDHTVNLLDTPC                       KDHTVNLLDTPC 80 90 40 150 FMNKYDREVRDS                     FMNKYDREVRDS                     FMNKYDREVRDS                       FMNKYDREVRDS                       FMNKYDREVRDS                         FMNKYDREVRDS                         FMNKYDREVRDS	EKLLLFSGAIQS  100  GHQDFSEDTYRV	SAGTVKGKKTGK 50  110  /LTAVDSALMVI           /LTAVDSALMVI 110  170  .KIRCAPVTWPI :          .QIRCAPVTWPI 170  230  CIQQLRDEIELV           CIQQLRDEIELV           CIQQLRDEIELV           CIQQLRDEIELV	120 DAAKGVEA         DAAKGVEA          DAAKGVEA          DAAKGVEA   20  180 GMGKNFKG          GMGKNFKG 180  240 QAASNEFN          QAASNEFN 240  300 PKFSGFIFK
m134.pep a134 m134.pep a134 m134.pep a134	MSQEILDO  IEKQRGIS  IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	VRRRRTFAII:	SHPDAGKTTLTE 20 30  80 90  KDHTVNLLDTPC 11         KDHTVNLLDTPC 80 90  40 150  FMNKYDREVRDS 11         FMNKYDREVRDS 40 150  LPHEFDIIKGIE 11         11        100 210  60 270  NNFGIQEILNSI 11         NNFGIQEILNSI 60 270	EKLLLFSGAIQS  40  100  SHQDFSEDTYR  SHQDFSEDTYR  100  160  SLELLDEVENII  SLELLDEVENII  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEQREPLE  CONPELEDREPLE  CONPELEQREPLE  CONPELEDREPLE  CONPELEQREPLE  CONPELEDREPLE  CONPELEDR	SAGTVKGKKTGK 50  110  /LTAVDSALMVI           /LTAVDSALMVI 110  170  LKIRCAPVTWPI 170  230  CIQQLRDEIELV           CIQQLRDEIELV           CIQQLRDEIELV 230  290  DATVRMVEPDEP            DATVRMVEPDEP 290	120 DAAKGVEA         DAAKGVEA          DAAKGVEA          DAAKGVEA   20  180 GMGKNFKG          GMGKNFKG          GMGKNFKG 180  240 QAASNEFN          QAASNEFN          CAASNEFN          CAASNEFN          CAASNEFN          CAASNEFN           CAASNEFN          CAASNEFN           CAASNEFN           CAASNEFN            CAASNEFN            CAASNEFN             CAASNEFN             CAASNEFN             CAASNEFN              CAASNEFN                 CAASNEFN
m134.pep a134 m134.pep a134 m134.pep a134	MSQEILDO  IEKQRGIS  IIIIIIII IEKQRGIS  1 QTIKLLNV IIIIIIII QTIKLLNV 1 VYHILNDE IIIIIIII VYHILNDE IIIIIIII LDEFLAGE IIIIIIII LDEFLAGE	70 8 5VASSVMQFDYI           5VASSVMQFDYI            5VASSVMQFDYI 70 8 .30 14 7CRLRDTPIVTI .30 14 7CRLRNTPIVTI .30 14 7CRLRNTPIVTI .30 26 EIYLFEAGGERI           EIYLFEAGGERI .90 26 ELTPVFFGSAIN           ELTPVFFGSAIN           ELTPVFFGSAIN           ELTPVFFGSAIN            ELTPVFFGSAIN            ELTPVFFGSAIN            ELTPVFFGSAIN            ELTPVFFGSAIN            ELTPVFFGSAIN            ELTPVFFGSAIN	SHPDAGKTTLTE 20 30  80 90  KDHTVNLLDTPC 11          KDHTVNLLDTPC 80 90  40 150  FMNKYDREVRDS 11          FMNKYDREVRDS 40 150  LPHEFDIKGIE 11           100 210  60 270  NNFGIQEILNSI 1            NNFGIQEILNSI 1          NNFGIQEILNSI 1	EKLLLFSGAIQS  40  100  GHQDFSEDTYRV  HILLIHI HILLIHI  GHQDFSEDTYRV  100  160  SLELLDEVENII  SLELLDEVENII  CONPELEQRFPLE  HILLIHIHI  DNPELEQRFPLE  HILLIHIHI  CONPELEQRFPLE  HILLIHIHIHI  HILLIHIHIHI  LIEWAPAPKPRI  CONPELEQRFPLE  HILLIHIHIHIHIHI  LIEWAPAPKPRI  CONPELEQRFPLE  HILLIHIHIHIHIHIHI  LIEWAPAPKPRI  HILLIHIHIHIHIHIHIHIHIHIHIHIHIHIHIHIHI	SAGTVKGKKTGK 50  110  VLTAVDSALMVI                      VLTAVDSALMVI 110  170  JKIRCAPVTWPI 170  230  EIQQLRDEIELV                      EIQQLRDEIELV   230  DATVRMVEPDEP                      DATVRMVEPDEP 290 350	TATSDWMD 60  120 DAAKGVEA         DAAKGVEA          DAAKGVEA 120  180 GMGKNFKG          GMGKNFKG 180  240 QAASNEFN          QAASNEFN           CAASNEFN           CAASNEFN           CAASNEFN           CAASNEFN            CAASNEFN            CAASNEFN            CAASNEFN             CAASNEFN             CAASNEFN             CAASNEFN              CAASNEFN               CAASNEFN
m134.pep a134 m134.pep a134 m134.pep a134	IEKQRGIS IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	TO SEVASSVMQFDYN SVASSVMQFDYN S	SHPDAGKTTLTE 20 30  80 90  KDHTVNLLDTPC 11          KDHTVNLLDTPC 80 90  40 150  FMNKYDREVRDS 11          FMNKYDREVRDS 40 150  00 210  LPHEFDIKGIE 11           LPHEFDIKGIE 11           100 270  NNFGIQEILNSI 11           NNFGIQEILNSI 11          NNFGIQEILNSI 100 270  20 330  CSGKFERGMKME	EKLLLFSGAIQS  40  100  GHQDFSEDTYR  HILLIHI HILLIH  GHQDFSEDTYR  100  160  SLELLDEVENII  HILLIHI HILLIH  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  ELLDEVENII  E	SAGTVKGKKTGK 50  110  VLTAVDSALMVI                      VLTAVDSALMVI   110  170  LKIRCAPVTWPI   170  230  EIQQLRDEIELV                    EIQQLRDEIELV   230  DATVRMVEPDEP                      DATVRMVEPDEP 290  350  SSVVTFMSHDRE	TATSDWMD 60  120 DAAKGVEA         DAAKGVEA          DAAKGVEA   20  180 GMGKNFKG   1      GMGKNFKG 180  240 QAASNEFN          QAASNEFN           CAASNEFN           CAASNEFN           CAASNEFN           CAASNEFN           CAASNEFN            CAASNEFN            CAASNEFN             CAASNEFN             CAASNEFN            CAASNEFN             CAASNEFN              CAASNEFN              CAASNEFN               CAASNEFN
m134.pep a134 m134.pep a134 m134.pep a134	MSQEILDO  IEKQRGIS            IEKQRGIS            QTIKLLNV            QTIKLLNV  1  VYHILNDE             VYHILNDE             LDEFLAGE             LDEFLAGE	TO SEVASSVMQFDYN SVASSVMQFDYN VASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQ	SHPDAGKTTLTE 20 30  80 90  KDHTVNLLDTPC 111111111111111111111111111111111111	EKLLLFSGAIQS  40  100  GHQDFSEDTYRV  SHQDFSEDTYRV  100  160  SLELLDEVENIN  SLELLDEVENIN  1111111111111111111111111111111111	SAGTVKGKKTGK 50  110  /LTAVDSALMVI                     /LTAVDSALMVI                   /LTAVDSALMVI 110  170  LKIRCAPVTWPI                   LQIRCAPVTWPI 170  230  CIQQLRDEIELV                       CIQQLRDEIELV                     CIQQLRDEIELV                     CIQQLRDEIELV                     CIQQLRDEIELV                     CIQQLRDEIELV                       CIQQLRDEIELV                       CIQQLRDEIELV                       CIQQLRDEIELV                       CIQQLRDEIELV                         CIQQLRDEIELV                         CIQQLRDEIELV                           CIQQLRDEIELV                           CIQQLRDEIELV                         CIQQLRDEIELV                           CIQQLRDEIELV                               CIQQLRDEIELV                                 CIQQLRDEIELV                                 CIQQLRDEIELV	TATSDWMD 60  120 DAAKGVEA          DAAKGVEA           DAAKGVEA 120  180 GMGKNFKG           GMGKNFKG 180  240 QAASNEFN          QAASNEFN           KFSGFIFK          KFSGFIFK 300  360 LVEEAYAG
m134.pep a134 m134.pep a134 m134.pep a134 m134.pep	MSQEILDO  IEKQRGIS            IEKQRGIS            QTIKLLNV            QTIKLLNV  1  VYHILNDE             VYHILNDE             LDEFLAGE             LDEFLAGE	TO SEVASSVMQFDYN SVASSVMQFDYN VASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFDN SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQFD SVASSVMQ	SHPDAGKTTLTE 20 30  80 90  KDHTVNLLDTPC 111111111111111111111111111111111111	EKLLLFSGAIQS  40  100  GHQDFSEDTYRV  SHQDFSEDTYRV  100  160  SLELLDEVENIN  SLELLDEVENIN  1111111111111111111111111111111111	SAGTVKGKKTGK 50  110  /LTAVDSALMVI                     /LTAVDSALMVI                   /LTAVDSALMVI 110  170  LKIRCAPVTWPI                   LQIRCAPVTWPI 170  230  CIQQLRDEIELV                       CIQQLRDEIELV                     CIQQLRDEIELV                     CIQQLRDEIELV                     CIQQLRDEIELV                     CIQQLRDEIELV                       CIQQLRDEIELV                       CIQQLRDEIELV                       CIQQLRDEIELV                       CIQQLRDEIELV                         CIQQLRDEIELV                         CIQQLRDEIELV                           CIQQLRDEIELV                           CIQQLRDEIELV                         CIQQLRDEIELV                           CIQQLRDEIELV                               CIQQLRDEIELV                                 CIQQLRDEIELV                                 CIQQLRDEIELV	TATSDWMD 60  120 DAAKGVEA          DAAKGVEA           DAAKGVEA 120  180 GMGKNFKG           GMGKNFKG 180  240 QAASNEFN          QAASNEFN           KFSGFIFK          KFSGFIFK 300  360 LVEEAYAG
m134.pep a134 m134.pep a134 m134.pep a134	MSQEILDO  IEKQRGIS  IIIIIIII IEKQRGIS  1 QTIKLLNV IIIIIIIII QTIKLLNV 1 VYHILNDE IIIIIIIII VYHILNDE IIIIIIIII LDEFLAGE IIIIIIIII LDEFLAGE IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	70 8 5VASSVMQFDYI                   5VASSVMQFDYI                   5VASSVMQFDYI                 70 8                   70 8                 70 8                   70 8                   70 10                   70 11                 70 12                 70 13                   70 14                 70 15                   70 16                   70 17                 70 18                   70 19                   70 19                   70 19                   70 19                   70 19                   70 19                   70 19                   70 19                   70 19                   70 19                   70 19                   70 19                   70 19                   70 19                   70 19                   70 19                   70 19                   70 19                   70 19                   70 19                   70 19                     70 19                     70 19                     70 19                     70 19                     70 19                     70 19                     70 19                     70 19                     70 19                     70 19                     70 19                     70 19                     70 19                     70 19                     70 19                     70 19                       70 19                       70 19                       70 19                       70 19                       70 19                       70 19                       70 19                       70 19                       70 19                       70 19                       70 19                           70 19                           70 19                               70 19	SHPDAGKTTLTE 20 30  80 90  KDHTVNLLDTPC 111111111111111111111111111111111111	EKLLLFSGAIQS  40  100  GHQDFSEDTYRV  HILLIHI HILLIHI  GHQDFSEDTYRV  100  160  SLELLDEVENII  HILLIHI HILLIHI  SLELLDEVENII  HILLIHI HILLIHI  DIPELEQRFPLE  HILLIHI HILLIHI  DIPELEQRFPLE  HILLIHI HILLIHI  DIPELEQRFPLE  HILLIHI HILLIHI  DIPELEQRFPLE  HILLIHI HILLIHI  DIPELEQRFPLE  HILLIHI HILLIHI  DIPELEQRFPLE  HILLIHI HILLIHI  DIPELEQRFPLE  HILLIHI HILLIHI  CHEMPAPAPKRI  HILLIHIHI  CHEMPAPAPKRI  HILLIHIHI  CHEMPAPAPKRI  HILLIHIHI  CHEMPAPAPKRI  HILLIHIHI  CHEMPAPAPKRI  HILLIHIHI  CHEMPAPAPKRI  CHEMPAPAPKRI  HILLIHIHIHI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPAPKRI  CHEMPAPKRI  CHE	SAGTVKGKKTGK 50  110  VLTAVDSALMVI                      VLTAVDSALMVI   110  170  LKIRCAPVTWPI   170  230  EIQQLRDEIELV                    EIQQLRDEIELV   230  DATVRMVEPDEP                      DATVRMVEPDEP 290  350  SSVVTFMSHDRE	TATSDWMD 60  120 DAAKGVEA          DAAKGVEA           DAAKGVEA 120  180 GMGKNFKG           GMGKNFKG 180  240 QAASNEFN          QAASNEFN           KFSGFIFK          KFSGFIFK 300  360 LVEEAYAG

m134.pep	370 DIIGIPNHGNIQIG	380 DSFSEGEQLA	390 AFTGIPFFAPE	400 ELFRSVRIKNI	410 PLKIKQLQKG	420 LQOLGEE
a134	DIIGIPNHGNIQIG	111111111	: ] [ ] [ ] [ ] [ ] [ ] [	 LFRSVRIKNI		 LQQLGEE
		300	390	400	410	420
m134.pep	430 GAVQVFKPMSGADL	440 ILGAVGVLQE	450 FEVVTSRLANE	460 YGVEAVEDSA	470	480
a134	GAVQVFKPMSGADL	1 1 1 1 1 1 1 1 1 1	11111	1111111111	11111111	
	430	440	450	460	470	480
m134.pep	490	500	510	520	530	
	AEFEKANAGNLAIDA		111111111	1111111111	1111111	
a134	AEFEKANAGNLAIDA 490	GGNLAYLAP 500	NRVNLGLTQE	RWPDIVFHET 520	REHSVKLX 530	

```
The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 517>:
```

```
ATGAAATACA AAAGAATCGT ATTTAAAGTC GGCACATCTT CGATTACCCG
  7
 51 TTCGGACGGC AGCCTCTCGC GCGGCAAAAT CCAAACCATC ACCCGCCAGC
101 TTGCCGCATT GCATCATGCG GGACACGAGC TGGTCTTGGT GTCTTCCGGC
     GCGGTTGCTG CAGGGTTCGG CGCGCTGGGT TTCAAAAAAC GTCCGGTCAA
201 AATCGCCGAC AAACAGGCTT CCGCCGCCGT CGGGCAGGGG CTGCTGATGG
251 AAGAATATAC GGCAAACCTG TCTTCAGACG GCATCGTGTC CGCACAAATC
301 CTGCTCAGCC GTGCCGACTT TGCCGACAAA CGCCGCTACC AAAATGCCGG
351 CGGCGCACTT TCCGTGCTGC TGCAACGCCG CGCGATTCCC ATCATCAATG
401 AAAACGACAC GGTTTCGGTT GAGGAGTTGA AAATCGGCGA CAACGACACA
451 TTGAGTGCGC AAGTGGCGGC GATGATACAG GCAGACCTCT TGGTGCTGCT
501 GACCGACATA GACGGTCTTT ACACCGGCAA CCCGAACAGC AATCCCGATG
551 CCGTACGGCT GGACAAAATC GAACACATCA ACCATGAAAT CATCGAAATG
    GCGGGCGGCT CGGGTTCGGC AAACGGCACG GGCGGTATGC TGACCAAAAT
    CAAAGCGGCA ACCATCGCCG CCGAATCCGG CGTACCGGTG TATATCTGTT
    CCTCACTCAA ACCCGATTCA TTGGCCGAAG CCGCCGAACA TCAGGCGGAC
701
751 GGCTCGTTTT TCGTcccCcg tgCCAAAGGT TTGCGGACAC AGAAGCAATG
801 GctggCGTTC TATTCcgaaa gcggGGgcag cgttTAtgtg gacgaaagtg
851 cggaacacgc tTtgtccgaa caagggaaag cctgCTGA
```

# This corresponds to the amino acid sequence <SEQ ID 518; ORF 135.ng>:

1 MKYKRIVFKV GTSSITRSDG SLSRGKIQTI TRQLAALHHA GHELVLVSSG
51 AVAAGFGALG FKKRPVKIAD KQASAAVGQG LLMEEYTANL SSDGIVSAQI
101 LLSRADFADK RRYQNAGGAL SVLLQRRAIP IINENDTVSV EELKIGDNDT
151 LSAQVAAMIQ ADLLVLLTDI DGLYTGNPNS NPDAVRLDKI EHINHEIIEM
201 AGGSGSANGT GGMLTKIKAA TIAAESGVPV YICSSLKPDS LAEAAEHQAD
251 GSFFVPRAKG LRTQKQWLAF YSESGGSVYV DESAEHALSE QGKAC*

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 519>:

1 ATGAAATACA AAAGAATCGT ATTTAAAGTC GGCACATCTT CGATTACCCA
51 TTCGGACGGC AGTCTCTCGC GCGCAAAAT CCAAACCATC ACCTGCCAGC
101 TTGCCGCATT GCATCATGCG GGACACGAGC TGGTCTTGGT GTCTTCCGGC
151 GCGGTTGCGG CAGGGTTCGG TGCGCTGGGT TTCAAAAAAC GTCCGGTCAA
201 AATCGCCGAC AAACAGGCTT CCGCCGCCGT CGGGCAGGG CTGCTGATGG
251 AAGAATATAC GGCAAACCTG TCTTCAGACG GCATCGTGTC CGCGCAAATC
301 CTGCTCAGCC GCGCGACTT TGCCGACAAA CGCCGCTACC AAAATGCCGG
351 CGGCGCACTT TCCGTGCTGC TGCAACGCCG CGCCGTCCC ATCATCAATG
401 AAAACGATAC GGTTTCGGTT GAGGAATTGA AAATCGGCGA CAACGACACA
451 TTGAGTGCCG AAGTGGCGGC GATGATACAG GCAGACCTCT TGGTGCTGCT

501	GACCGACATA	GACGGTCTTT	ACACGGGCAA	CCCGAACAGC	AATCCCGATG
551	CCGTACGGCT	GGACAAAATC	GAACACATCA	ACCATGAAAT	CATCGAAATG
601	GCGGGCGGCT	CGGGTTCGGC	AAACGGCACG	GGCGGTATGC	TGACCAAAAT
651	CAAAGCGGCA	ACCATCGCCG	CCGAATCCGG	CGTACCGGTG	TATATCTGTT
701	CCTCGCTCAA	ACCCGATGCA	CTTGCCGAAG	CTGCCGAACA	TCAGGCGGAC
751	GGCTCGTTTT	TCGTCCCCCG	TGCCAAAGGT	TTGCGGACGC	AGAAGCAATG
801	GCTGGCGTTC	TATTCCGAAA	GCCGGGGCAG	CGTTTATGTG	GACGAAGGTG
851	CGGAACACGC	TTTGTCCGAA	CAGGGGAAAA	GCCTGCTGAT	GTCGGGCATT
901	GCCGGAATCG	AAGGGCATTT	TTCCCGTATG	GACACCGTAA	CCGTGTACAG
951	CAAGGCAACC	AAACAGCCCC	TGGGCAAAGG	GCGCGTCCTG	TTCGGCTCTG
1001	CCGCCGCCGA	AGACCTGCTC	AAATCGCGTA	AGGCGAAAGG	CGTGTTCATC
1051	CATCGGGACG	ACTGGATTTC	CATCACGCCC	GAAATACGCC	TGCTTCTGAC
1101	CGAATTTTAG				

This corresponds to the amino acid sequence <SEQ ID 520; ORF 135>:

m135.pep

1 MKYKRIVFKV GTSSITHSDG SLSRGKIQTI TCQLAALHHA GHE<u>LVLVSSG</u>
51 <u>AVAAGFGALG</u> FKKRPVKIAD KQASAAVGQG LLMEEYTANL SSDGIVSAQI
101 LLSRADFADK RRYQNAGGAL SVLLQRRAVP IINENDTVSV EELKIGDNDT
151 <u>LSAQVAAMIQ</u> ADLLVLLTDI DGLYTGNPNS NPDAVRLDKI EHINHEIIEM
201 AGGSGSANGT GGMLTKIKAA TIAAESGVPV YICSSLKPDA LAEAAEHQAD
251 GSFFVPRAKG LRTQKQWLAF YSESRGSVYV DEGAEHALSE QGKSLLMSGI
301 AGIEGHFSRM DTVTVYSKAT KQPLGKGRVL FGSAAAEDLL KSRKAKGVFI
351 HRDDWISITP EIRLLTEF*

Computer analysis of this amino acid sequence gave the following results:

10

Homology with a predicted ORF from N. gonorrhoeae

ORF 135 shows 97.6% identity over a 294 aa overlap with a predicted ORF (ORF 135.ng) from N. gonorrhoeae:

20

3.0

40

50

60

m135/g135

	10	20	30	40	50	60
m135.pep	MKYKRIVFKVGTSSI	THSDGSLSRG	KIQTITCQLA	ALHHAGHEL'	VLVSSGAVAA	GFGALG
		1:1111111	111111111		[]]]]	
g135	MKYKRIVFKVGTSSI	TRSDGSLSRG	KIQTITRQLA	ALHHAGHEL	VLVSSGAVAA	GFGALG
	10	20	30	40	50	60
	70	80	90	100	110	120
m135.pep	FKKRPVKIADKQASA	AVGQGLLMEE	YTANLSSDGI	VSAQILLSR	ADFADKRRYO	NAGGAL
		1111111111				
g135	FKKRPVKIADKQASA	AVGQGLLMEE	YTANLSSDGI	VSAQILLSR	ADFADKRRYO	NAGGAL
	70	80	90	100	110	120
	130	140	150	160	170	180
m135.pep	SVLLQRRAVPIINEN	DTVSVEELKI	GDNDTLSAQV	AAMIQADLL	VLLTDIDGLY	TGNPNS
-		11111111111		HHIIIII		111111
g135	SVLLORRAIPIINEN	DTVSVEELKI	GDNDTLSAOV	AAMIOADLL	VLLTDTDGLY	
3	130	140	150	160	170	180
						200
	190	200	210	220	230	240
m135.pep	NPDAVRLDKIEHINH	EIIEMAGGSG	SANGTGGMLT	KIKAATTAA		
				111111111		11111.
g135	NPDAVRLDKIEHINH	ETTEMAGGSG	SANCTCCMI.T		FSGVDVVTCS	יווווי
5200	190	200	210	220	230	240
	130	200	210	220	230	240
	250	260	270	280	290	300
m135.pep	LAEAAEHQADGSFFV					
g135	11111111111111111111111111111111111111			,,,,,,,		
9133	LAEAAEHQADGSFFV 250					CX
	250	260	270	280	290	
	310	320	330	240	250	
	310	320	330	340	350	360

WO 99/57280



m135.pep AGIEGHFSRMDTVTVYSKATKQPLGKGRVLFGSAAAEDLLKSRKAKGVFIHRDDWISITP

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 521>:

```
ATGAAATACA AAAGAATCGT ATTTAAAGTC GGCACATCTT CGATTACCCA
       TTCGGACGGC AGTCTCTCGC GCGGCAAAAT CCAAACCATC ACCCGCCAGC
   51
       TTGCCGCATT GCATCATGCG GGACACGAGC TGGTCTTGGT GTCTTCCGGC
  101
       GCGGTTGCGG CAGGGTTCGG TGCGCTGGGT TTCAAAAAAC GTCCGGTCAA
  151
 201 AATCGCCGAC AAACAGGCTT CCGCCGCCGT CGGGCAGGGG CTGCTGATGG
 251 AAGAATATAC GGCAAAACCTG TCTTCAGACG GCATCGTGTC CGCACAAATC
301 CTGCTCAGCC GCGCCGACTT TGCCGACAAA CGCCGCTACC AAAATGCCGG
      CGGCGCACTT TCCGTGCTGC TGCAACGCCG CGCCGTCCCC ATCATCAATG
      AAAACGATAC GGTTTCGGTT GAGGAATTGA AAATCGGCGA CAACGACACA
 401
      TTGAGTGCGC AAGTGGCGGC GATGATACAG GCAGACCTCT TGGTGCTGCT
 451
      GACCGACATA GACGGTCTTT ACACCGGCAA CCCGAACAGC AATCCCGATG
 501
 551 CCGTACGGCT GGACAAAATC GAACACATCA ACCATGAAAT CATCGAAATG
      GCGGGCGGCT CGGGTTCGGC AAACGGCACA GGCGGTATGC TGACTAAAAT
 601
      CAAAGCGGCG ACGATTGCGA CCGAGTCCGG CGTACCGGTC TATATCTGTT
 651
 701 CCTCGCTCAA ACCCGATGCA CTTGCCGAAG CGGCAGATAA TCAGGCGGAC
 751 GGCTCGTTTT TCGTCCCCCG TGCCAAAGGT TTGCGGACGC AGAAGCAATG
 801 GCTGGCGTTC TATTCCGAAA GCAGGGGCGG CGTTTATGTG GACGAAGGTG
     CGGAACACGC TTTGTCCGAA CAGGGAAAAA GCCTGCTGAT GTCGGGCATT
 851
 901
      GCCGGAATCG AAGGGCATTT TTCCCGTATG GACACCGTAA CCGTGTACAG
 951 CAAGGCAACC AAACAGCCTT TGGGCAAAGG GCGAGTCCTG TTCGGCTCTG
1001 CCGCCGCCGA AGACCTGCTC AAATTGCGTA AGGCGAAAGG CGTGTTCATC
1051 CATCGGGACG ACTGGATTTC CATCACGCCC GAAATACGCC TGCTTCTGAC
1101 CGAATTTTAG
```

## This corresponds to the amino acid sequence <SEQ ID 522; ORF 135.a>:

a135.pep

1 MKYKRIVFKV GTSSITHSDG SLSRGKIQTI TRQLAALHHA GHELVLVSSG
51 AVAAGFGALG FKKRPVKIAD KQASAAVGQG LLMEEYTANL SSDGIVSAQI
101 LLSRADFADK RRYQNAGGAL SVLLQRRAVP IINENDTVSV EELKIGDNDT
151 LSAQVAAMIQ ADLLVLLTDI DGLYTGNPNS NPDAVRLDKI EHINHEIIEM
201 AGGSGSANGT GGMLTKIKAA TIATESGVPV YICSSLKPDA LAEAADNQAD
251 GSFFVPRAKG LRTQKQWLAF YSESRGGVYV DEGAEHALSE QGKSLLMSGI
301 AGIEGHFSRM DTVTVYSKAT KQPLGKGRVL FGSAAAEDLL KLRKAKGVFI

## m135/a135 98.4% identity in 369 aa overlap

m135.pep	10 MKYKRIVFKVGTS !!!!!!!!!!!	20 SITHSDGSLSI 	30 RGKIQTITCQ	40 LAALHHAGHEI	50 VLVSSGAVA	60 AGFGALG
a135	MKYKRIVFKVGTS	SITHSDGSLS				111111
	10	20	30	40		
			30	40	50	60
-125	70	80	90	100	110	100
m135.pep	FKKRPVKIADKQA	SAAVGQGLLME	EEYTANLSSD	GIVSAOILLSR	AULTIO	T20
a135						
a133	FKKRPVKIADKQAS	SAAVGQGLLME	EYTANLSSDO	GIVSAQILLSR	ADFADKRRY	וווווו
	70	80	90	100	110	120
	130	140	150			
m135.pep		TAO TAO	150	160	170	180
	SVLLQRRAVPIINE	IIIIIIIII	TGDNDTLSAC	SAWWIOWDIT	VLLTDIDGLY	TGNPNS
a135	SVLLQRRAVPIINE	NDTVSVEELK	TGDNDTTCAC			111111
	130	140	150	160		TGNPNS
			100	160	170	180
-125	190	200	210	220	230	
m135.pep	NPDAVRLDKIEHIN	HEIIEMAGGS	GSANGTGGML	TKIKAATTAA	23U SSGVDVVTCC	240
a135						
4133	NPDAVRLDKIEHIN 190	HEIIEMAGGS(	GSANGTGGML	TKIKAATIAT	SGVPVYTCS	ארסאד קרווון
	190	200	210	220	230	240
					•	-40

270

200

200

250

	25	50 20	60 27	0 280	290	300
m135.pep	LAEAAEHQA	DGSFFVPRA	KGLRTQKQWLA	FYSESRGSVYV	DEGAEHALSEQ	SKSLLMSGI
	::			111111111111		11111111
a135	LAEAADNQA	DGSFFVPRA	KGLRTQKQWLA	FYSESRGGVYV	DEGAEHALSEQ	SKSLLMSGI
	25	50 26	60 27	0 280	290	300
	31		20 33			360
m135.pep					KSRKAKGVFIHI	
a135					KLRKAKGVFIHI	
	31	.0 32	20 33	340	350	360
405	37					
m135.pep	EIRLLTEF					
405	11111111					
.a135	EIRLLLTEF					
	37	0				
The fellowing a	antial DNIA and		. : 4 - 4 : 6 - 4 : -	• λ77-	<ceo< td=""><td>TD 5005</td></ceo<>	TD 5005
The following pa	antial DNA sec	quence was	i identified fi	n IV. <b>gonorrn</b>	oeae <seq< td=""><td>ID 523&gt;:</td></seq<>	ID 523>:
g136. <b>seq</b>						
1	ATGGAAATCC G					
51	AAACGCTtca a					
101	CACGGACAGG G					
151	TTGCGGTTTG T					
201	cgTAAGGCag t					
251	AGgcggataa c					
301	cacgGGGTAA A					
351	AATCGGTCGG C					
401	TAAAAATCGC T	GCCGCCCTC	GCCGTCGTCG	TAGAGCCACA	AATCGGGCAG	
451	CTTTTTATCC G	ACATCGCGG	CGGTTGTTTC	CATCGCCATT	GCCAAAACCA	
501	GCCGTTCGAT T	TCGGAACGT	TCGGCGGCGG	TAAATTGCGA	TTCGTCGCCC	
551	AACACTTCGG G	CAGCCAGTC	GAGCGGTGCC	AATTTGTCCG	GCCCGCTCAA	
601	CAGCGCCGTC A	TAAAACCTT	GAACCTCGTC	GCAACGCATC	GTGTTGCCTT	
651	GTTCGCTTTT G	GCATCCAAT	AA			
This correspond	s to the amino	acid seque	nce <seo ii<="" td=""><td>D 524: ORF</td><td>136.ng&gt;:</td><td></td></seo>	D 524: ORF	136.ng>:	
g136.pep				, 0.44		
9150.pcp 1	MEIRFQTAFL R	T.VOMKTNA C	TI.TATRI.VFP	ΔΔΔΔΡΤΩΤΩΡ	AGEFDFDADG	
51	LRFVDDRLPV A					
21	ARE VOURLEV M	**P * COV * VO	TOURTRUDAL	GETÖVDIVAL	TIE A AIMMAICH	

201 QRRHKTLNLV ATHRVALFAF GIQ*
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 525>:

101 HGVKQLFKRF IIGGFKPIGR HNVQTVKIGV APSVKIAAAL AVVVEPQIGQ 151 LFIRHRGGCF HRHCQNQPFD FGTFGGGKLR FVAQHFGQPV ERCQFVRPAQ

m136.seq ATGGAAACAA ACGCTTCAAT TCTTACCGCA ACACGCCTTG TATTTTCTGC 51 CGCTGCCGCA CGGACAGGGA TCGTTCCTGC CTGTTTTTTC GCCTTCCCTG 101 CGGACGGTTT GCGGTTTGTT GATGACTGCC TGCCAGTAGC GGTAGATATC 151 CGCCAATGCA TAAGGCAACT CGGATTCCAG TTCCGCCAGC TCGCCTTCTG 201 TGAATTGCAG ACGGATAGCG CCGTTTTCCT CTTCGTCGTA AATACCGCCC 251 AATGCCATGA TGGGATAAAA CAACTCTTCA AACGCTTCAT CATCGACGGC TTCAAACCAA TCGGTCGGCA CAATATCCAA ACCGTAAAGA TAAGCATTGC 351 ACCATGTGTA AAAATCGCTG CCGCCGTCTT CGTTTTCATA CAGCCACAAA 401 TCGGGCAGTT TTTTATCCGA CATCGCGGCG GTTGTTTCCA TCGCCATTGC 451 CAAAACCAGC CGTTCGATTT CGGAACGTTC GGCGGCGGTA AATTGCGATT 501 CGTCGCCCAA CACTTCGGGC AGCCAGTCGA GCGGTGTCAA TTTGTCCGGC 551 CCGCTCAACA GCGCCGTCAT AAAACCTTGA ACCTCGTCGC AACGCATCGT 601 GTTGCCTTGT TCGCTTTTGG CATCCAACAA TTCGCTCAAC CGCCGTTTGG 651 ATGCTTCGGT AAATTTTCGG GAATCCATCA TTTTCCTTTT CAAATGGGTT TTGCGCCCTA TTATCGCCGC AATGCCGTCT GA

This corresponds to the amino acid sequence <SEQ ID 526; ORF 136>:

```
m136.pep
                METNASILTA TRLVFSAAAA RTGIVPACFF AFPADGLRFV DDCLPVAVDI
             1
                RQCIRQLGFQ FRQLAFCELQ TDSAVFLFVV NTAQCHDGIK QLFKRFIIDG
            51
                FKPIGRHNIQ TVKISIAPCV KIAAAVFVFI QPQIGQFFIR HRGGCFHRHC
                QNQPFDFGTF GGGKLRFVAQ HFGQPVERCQ FVRPAQQRRH KTLNLVATHR
                VALFAFGIQQ FAQPPFGCFG KFSGIHHFPF QMGFAPYYRR NAV*
 Computer analysis of this amino acid sequence gave the following results:
 Homology with a predicted ORF from N. gonorrhoeae
 ORF 136 shows 85.6% identity over a 209 aa overlap with a predicted ORF (ORF 136.ng)
 from N. gonorrhoeae:
      m136/g136
                                       10
                                                20
                                                          30
      m136.pep
                               METNASILTATRLVFSAAAARTGIVPACFFAFPADGLRFVDDCLPV
                               {\tt MEIRFOTAFLRLVQMKTNASILTATRLVFPAAAARTGIVPAGFFPFPADGLRFVDDRLPV}
      g136
                          10
                                   20
                                            30
                                                      40
                                                               50
                             60
                                       70
                                                          90
                  {\tt AVDIRQCIRQLGFQFRQLAFCELQTDSAVFLFVVNTAQCHDGIKQLFKRFIIDGFKPIGR}
      m136.pep
                  AVDVCQRVRQFGRKFRQLAFGELQADNAVFLFVVNAAHCHHGVKQLFKRFIIGGFKPIGR
      g136
                          70
                                   80
                                            90
                                                    100
                                                                        120
                   110
                            120
                                     130
                                               140
                                                        150
                                                                  160
     m136.pep
                  \verb| HNIQTVKISIAPCVKIAAAVFVFIQPQIGQFFIRHRGGCFHRHCQNQPFDFGTFGGGKLR|
                  HNVQTVKIGVAPSVKIAAALAVVVEPQIGQLFIRHRGGCFHRHCQNQPFDFGTFGGGKLR
     9136
                        130
                                 140
                                           150
                                                    160
                                                              170
                                                                       180
                            180
                                     190
                                              200
                                                        210
                                                                  220
                 {\tt FVAQHFGQPVERCQFVRPAQQRRHKTLNLVATHRVALFAFGIQQFAQPPFGCFGKFSGIH}
     m136.pep
                  q136
                 FVAQHFGQPVERCQFVRPAQQRRHKTLNLVATHRVALFAFGIQX
                        190
                                 200
                                           210
                  230
                            240
     m136.pep
                 HFPFQMGFAPYYRRNAVX
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 527>:
     a136.seq
              ATGGAAACAA ACGCTTCAAT TCTTACCGCA ACACGCCTTG TATTTTCTGC
              CGCTGCCGCA CGGACAGGGA TCGTTCCTGC CTGTTTTTTC GCCTTCCCTG
          51
             CGGACGGTTT GCGGCTTGTT GATGACCGCC TGCCAGTAGC GGTAGATATC
         101
         151 CGCCAATGCA TAAGGCAACT CGGATTCCAG TTCCGCCAGC TCGCCTTCTG
             TGAATTGCAG ACGGATAGTG CCGTTGTCCT CTTCGTCGTA AATACCGCCC
         201
             AATGCCATGA TGGGATAAAA CAACTCTTCA AACGCTTCAT CATCGACGGC
         251
             TTCAAACCAA TCGGTCGGCA CAATATCCAA ACCGTAAAGA TAAGCATTGC
         301
             ACCATGTGTA AAAATCGCTG CCGCCGTCTT CGTTTTCATA CAGCCACAAA
         351
             TCGGGCAGTT TTTTATCCGA CATCGCGGCG GTTGTTTCCA TCGCCATTGC
         401
             CAAAACCAGC CGTTCGATTT CGGAACGTTC GGCGGCGGTA AATTGCGATT
         451
             CGTCGCCCAA CACTTCGGGC AGCCAGTCGA GCGGTGTCAA TTTGTCCGGC
         501
             CCGCTCAACA GCGCCGTCAT AAAACCTTGA ACCTCGTCGC AACGCATCGT
         551
             GTTGCCTTGT TCGCTTTTGG CATCCAACAA TTCGCTCAAC CGCCGTTTGG
         601
         651 ATGCTTCGGT AAATTTTCGG GAATCCATCA TTTTCCTTTT CCAATGGGTT
             TTGCGCCCTA TTATAGTGGA TTAAATTTAA ATCAGGACAA GGCGACGAAG
             CCGCAGACAG TACAAATAGT ACGGCAAGGC GAGGCAACGC CGTACTGGTT
         751
```

This corresponds to the amino acid sequence <SEQ ID 528; ORF 136.a>: a136.pep

TAAATTTAAT CCACTATATC GCCGCAATGC CGTCTGA

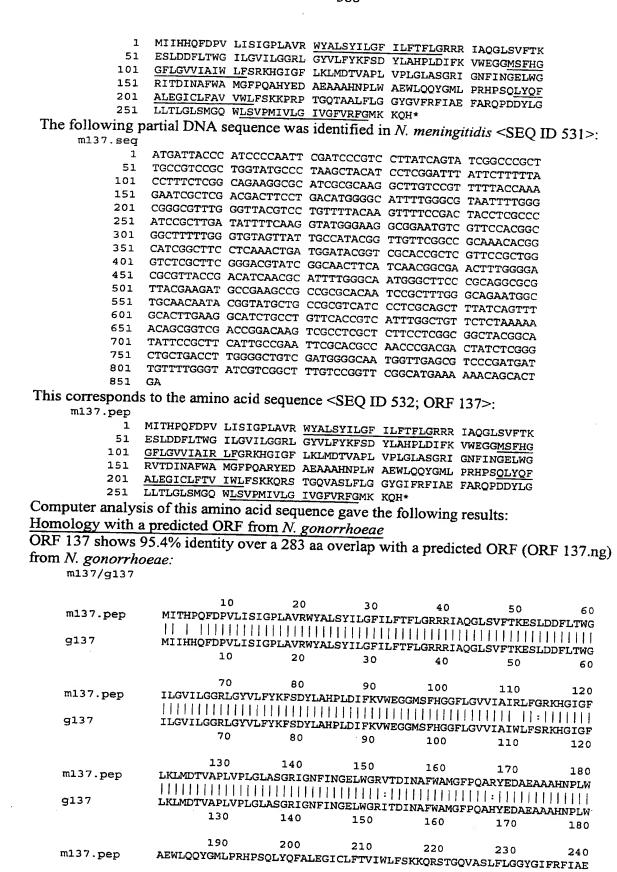
1 51 101 151 201 251	METNASILTA TRLVFS: RQCIRQLGFQ FRQLAFG FKPIGRHNIQ TVKISII QNQPFDFGTF GGGKLRI VALFAFGIQQ FAQPPFG PQTVQIVRQG EATPYWI	CELO TDSAV APCV KIAAJ FVAQ HFGQI GCFG KFSGI FKFN PLYRI	VVLFVV NTAC AVFVFI QPQI PVERCQ FVRF IHHFPF PMGF	CHDGIK QLE GQFFIR HRG AQQRRH KTI	FKRFIIDG GGCFHRHC NLVATHR	
m136/a136 98	3.3% identity in 238 a	aa overlap				
	10	20	30	40	50	60
m136.pep	METNASILTATRLV	FSAAAARTG:	IVPACFFAFPA	DGLRFVDDCI	PVAVDIRQCI	RQLGFQ
-136				1111:111		
a136	METNASILTATRLVI	20	30	DGLKLVDDRI 40	PVAVDIRQCI 50	
	10	. 20	30	40	30	60
	70	. 80	90	100	110	120
m136.pep	FRQLAFCELQTDSAV	/FLFVVNTAC	CHDGIKQLFK	RFIIDGFKPI		
				111111111	1111111111	HILLI
a136	FRQLAFCELQTDSAV				GRHNIQTVKI	SIAPCV
	70	80	90	100	110	120
	130	140	1.50	1.60	450	
m136.pep	KIAAAVFVFIQPQIO		150	160	170	180
mrso.beb		50111111100		IIIIIIIII	LILLLLLLL	PVERCQ
a136	KIAAAVFVFIQPQIO	OFFIRHRGO	CEHRHCONOP	FDFGTFGGGK	LREVAOHEGO	PVERCO
	130	140	150	160	170	180
	190	200	210	220	230	240
m136.pep	FVRPAQQRRHKTLNI	LVATHRVALE	TAFGIQQFAQP	PFGCFGKFSG	IHHFPFQMGF	APYYRR
a136						
a130	FVRPAQQRRHKTLNI 190	200	AFGIQQFAQF 210	220	230	APYYSG 240
	230	200	210	220	230	240
m136.pep	XVAN					
a136	LNLNQDKATKPQTVQ	-		RNAVX		
	250	260	270			

# The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 529>: g137.seq

```
ATGATTATCC ATCACcaaTT CGATCCCGTC CTCATCAGTA TCGGCCCGCT
 1
    TGCCGTCCGC TGGTATGCCT TAAGCTACAT CCTCGGATTT ATTCTTTTTA
 51
    CCTTTCTCGG CAGAAGGCGC ATCGCGCAAG GCTTGTCCGT TTTTACCAAA
151
    GAATCGCTCG ACGACTTCCT GACATGGGGC ATTTTGGGCG TGATTTTGGG
201 CGGACGCTTG GGCTATGTCC TGTTTTACAA ATTCTCCGAC TACCTCGCCC
251 ATCCGCTTGA TATTTTCAAG GTATGGGAAG GCGGAATGTC GTTCCACGGC
301 GGCTTTTTGG GTGTAGTTAT TGCCATATGG TTGTTCAGCC GCAAGCACGG
351 CATCGGCTTC CTCAAACTGA TGGACACGGT CGCGCCGCTC GTTCCGCTGG
401 GTCTCGCTTC GGGACGTATC GGCAACTTTA TCAACGGCGA ACTTTGGGGA
451 CGCATTACCG ACATTAACGC ATTTTGGGCA ATGGGCTTCC CGCAAGCGCA
501 TTACGAAGAT GCCGAAGCCG CCGCGCACAA TCCGCTTTGG GCAGAATGGC
551 TGCAACAATA CGGTATGCTG CCGCGTCATC CCTCGCAGCT TTATCAGTTT
601 GCCCTTGAAG GCATCTGCCT GTTCGCCGTC GTTTGGCTGT TTTCCAAAAA
    ACCGCGCCCG ACCGGGCAGA CTGCCGCGCT TTTTCTCGGC GGCTACGGCG
    TGTTCCGCTT TATTGCCGAA TTTGCGCGCC AACCCGACGA CTATCTCGGG
    CTGCTGACCT TGGGGCTGTC GATGGGGCAA TGGTTGAGCG TCCCGATGAT
    TGTTTTGGGT ATCGTCGGCT TTGTCCGGTT CGGCATGAAA AAACAGCACT
801
```

This corresponds to the amino acid sequence <SEQ ID 530; ORF 137.ng>: g137.pep

851



g137	
m137.pep	250 260 270 280 FARQPDDYLGLLTLGLSMGQWLSVPMIVLGIVGFVRFGMKKQHX
mrs / . pep	
g137	FARQPDDYLGLLTLGLSMGQWLSVPMIVLGIVGFVRFGMKKQHX 250 260 270 280
• • • • • • • • • • • • • • • • • • • •	partial DNA sequence was identified in N. meningitidis <seq 533="" id="">:</seq>
a137.seq 1	ATGATTACCC ATCCCCAATT CGACCCCGTC CTTATCAGTA TCGGCCCGCT
51	TGCCGTCCGC TGGTATGCCC TAAGCTACAT CCTCGGATTT ATTCTTTTA
101	CCTTTCTCGG CAGAAGGCGC ATCGCGCAAG GCTTGTCCGT TTTTACCAAA
151	GAATCGCTCG ACGACTTCCT GACATGGGGC ATTTTGGGCG TAATTTTGGG
201 251	CGGGCGTTTG GGTTACGTCC TGTTTTACAA GTTTTCCGAC TACCTCGCCC ATCCGCTTGA TATTTTCAAG GTATGGGAAG GCGGAATGTC GTTCCACGGC
301	GGCTTTTTGG GTGTAGTTAT TGCCATATGG TTGTTCGGTC GCAAACACGG
351	CATCGGCTTC CTCAAACTGA TGGACACGGT CGCACCGCTC GTTCCACTGG
401	GTCTCGCTTC GGGACGTATC GGCAACTTCA TCAACGGCGA ACTTTGGGGA
451	CGCGTTACCG ACATCAACGC ATTTTGGGCA ATGGGCTTCC CGCAGGCGCG TTACGAAGAC CTCGAAGCCG CCGCGCACAA TCCGCTTTGG GCAGAATGGC
501 551	TGCAACAATA CGGTATGCTG CCGCGTCATC CCTCGCAGCT TTATCAGTTT
601	GCACTTGAAG GCATCTGCCT GTTCGCCGTC GTTTGGCTGT TCTCTAAAAA
651	ACAGCGGCCG ACCGGACAAG TCGCCTCACT CTTCCTCGGC GGCTACGGCA
701	TATTCCGCTT CATTGCCGAA TTTGCACGCC AACCCGACGA CTATCTCGGG
751 801	CTGCTGACCT TGGGGCTGTC GATGGGGCAA TGGTTGAGCG TCCCGATGAT TGTTTTGGGT ATCGTCGGCT TTGTCCGGTT CGGCATGAAA AAACAGCACT
851	GA
<del>-</del>	Is to the amino acid sequence <seq 137.a="" 534;="" id="" orf="">:</seq>
a137.pep	WITH THE PROPERTY AND THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROP
1 51	MITHPQFDPV LISIGPLAVR <u>WYALSYILGF ILFTFLG</u> RRR IAQGLSVFTK ESLDDFLTWG ILGVILGGRL GYVLFYKFSD YLAHPLDIFK VWEGGMSFHG
101	GFLGVVIAIW LFGRKHGIGF LKLMDTVAPL VPLGLASGRI GNFINGELWG
151	RVTDINAFWA MGFPQARYED LEAAAHNPLW AEWLQQYGML PRHPSQLYQF
201	ALEGICLFAV VWLFSKKQRP TGQVASLFLG GYGIFRFIAE FARQPDDYLG
251	LLTLGLSMGQ WLSVPMIVLG IVGFVRFGMK KQH*
m137/a137 98	3.2% identity in 283 aa overlap
m137.pep	10 20 30 40 50 60 MITHPQFDPVLISIGPLAVRWYALSYILGFILFTFLGRRRIAQGLSVFTKESLDDFLTWG
a137	MITHPQFDPVLISIGPLAVRWYALSYILGFILFTFLGRRRIAQGLSVFTKESLDDFLTWG
	10 20 30 40 50 60
	70 80 90 100 110 120
m137.pep	ILGVILGGRLGYVLFYKFSDYLAHPLDIFKVWEGGMSFHGGFLGVVIAIRLFGRKHGIGF
a137	
a13/	70 80 90 100 110 120
	130 140 150 160 170 180
m137.pep	LKLMDTVAPLVPLGLASGRIGNFINGELWGRVTDINAFWAMGFPQARYEDAEAAAHNPLW
a137	LKLMDTVAPLVPLGLASGRIGNFINGELWGRVTDINAFWAMGFPQARYEDLEAAAHNPLW
	130 140 150 160 170 180
	190 200 210 220 230 240
m137.pep	190 200 210 220 230 240 AEWLQQYGMLPRHPSQLYQFALEGICLFTVIWLFSKKQRSTGQVASLFLGGYGIFRFIAE
• •	
a137	AEWLQQYGMLPRHPSQLYQFALEGICLFAVVWLFSKKQRPTGQVASLFLGGYGIFRFIAE

BNSDOCID: <WO___9957280A2_I_>

0.00

190 200 210 220 230 240 250 260 270 280 FARQPDDYLGLLTLGLSMGQWLSVPMIVLGIVGFVRFGMKKQHX FARQPDDYLGLLTLGLSMGQWLSVPMIVLGIVGFVRFGMKKQHX a137 250 260 270

```
The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 535>:
```

```
ATGGAGTTTG AAAACATTAT TTCCGCCGCc gaCAAGGCGC GTATCCTTGC
  1
     CGAAGCACTG CCTTACAtcc gccgGTTTTC CGGTTCGGTC GCCGTCATCA
 51
101 AGTATGGCGG CAACGCGATG ACCGAACCTG CCTTGAAAGA AGGGTTTGCC
151 CGCGATGTCG TGCTGCTGAA GCTGGTCGGC ATTCATCCCG TCATCGTTCA
201 CGGCGGCGGG CCGCAGATCA ATGCGATGCT TGAAAAAGTC GGCAAAAAGG
251 GCGAATTTGT CCAAGGAATG CGCGTTACCG ACAAAGAGAC GATGGATATT
301 GTCGAAATGG TATTGGGCGG GCACGTCAAC AAGGAAATCG TGTCGATGAT
351 TAACACATAT GGAGGGCACG CGGTCGGCGT GAGCGGCGC GACGACCATT
     TCATTAAGGC GAAGAAACTT TTGGTCGATA CGCCCGAACA GAATAGCGTG
401
     GACATCGGAC AGGTCGGTAC GGTGGAAAGC ATCGATACCG GTTTGGTTAA
451
     AGGGCTGATA GAACGCGGCT GCATTCCCGT CGTCGCCCCC GTCGGCGTAG
     GTGAAAAAGG CGAAGCGTTC AACATCAACG CCGATTTGGT GGCAGGCAAA
551
601 TTGGCGGAAG AATTGAACGC CGAAAAACTC TTGATGATGA CGAAtatcgc
651 CGGTGTGATG GACAAAACGG GCAATCTGCT GACCAAACTC ACGCCGAAAC
701 GGATTGATGG GCTGATTGCC GACGGCACGC TGTATGGCGG TATGCTGCCG
751 AAAATCGCTT CTGCGGTCGA AGCCGCGtc aACGGTGTGA AAGCCACGCA
801 CATCATCGAC GGCAGGTTGC CCAACGCGCT TTTGCTGGAA ATCTTTACCG
851 ATGCCGGTAT CGGGTCGATG ATTTTAGGCA GAGGGGAAGA TGCCTGA
```

# This corresponds to the amino acid sequence <SEQ ID 536; ORF 138.ng>:

```
1 MEFENIISAA DKARILAEAL PYIRRFSGSV AVIKYGGNAM TEPALKEGFA
51 RD<u>VVLLKLVG IHPVIVHGG</u>G PQINAMLEKV GKKGEFVQGM RVTDKETMDI
101 VEMVLGGHVN KEIVSMINTY GGHAVGVSGR DDHFIKAKKL LVDTPEQNSV
151 DIGQVGTVES IDTGLVKGLI ERGCIPVVAP VGVGEKGEAF NINADLVAGK
```

201 LAEELNAEKL LMMTNIAGVM DKTGNLLTKL TPKRIDGLIA DGTLYGGMLP 251 KIASAVEAAV NGVKATHIID GRLPN<u>ALLLE IFTDAGIGSM IL</u>GRGEDA*

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 537>:

```
ATGGAGTCTG AAAACATTAT TTCCGCCGCC GACAAGGCGC GTATCCTTGC
  1
    CGAAGCGCTG CCTTACATCC GCCGGTTTTC CGGTTCGGTC GCCGTCATCA
 51
    AATACGGCGG CAACGCGATG ACCGAACCTG CCTTGAAAGA AGGGTTTGCC
101
     CGCGATGTCG TGCTGCTGAA GCTGGTCGGC ATTCATCCCG TCATCGTTCA
201 CGGCGGCGGG CCGCAGATCA ATGCGATGCT TGAAAAAGTC GGCAAAAAGG
251 GTGAGTTTGT CCAAGGAATG CGCGTTACCG ACAAAGAGGC GATGGATATT
301 GTCGAAATGG TGTTGGGCGG GCATGTCAAT AAAGAAATCG TGTCGATGAT
351 TAACACATAT GGCGGACACG CGGTCGGCGT AAGCGGACGC GACGACCATT
401 TCATTAAGGC GAAGAAACTT TTGATCGATA CGCCCGAACA GAATGGCGTG
451 GACATCGGAC AGGTCGGTAC GGTGGAAAGC ATCGATACCG GTTTGGTTAA
501 AGGGCTGATA GAACGTGGCT GCATTCCCGT CGTCGCCCCC GTCGGCGTAG
551 GTGAAAAAGG CGAAGCGTTC AACATCAACG CCGATTTGGT AGCAGGCAAA
    TTGGCGGAAG AATTGAACGC CGAAAAACTC TTGATGATGA CGAATATCGC
    CGGTGTGATG GACAAAACGG GCAATCTGCT GACCAAACTC ACGCCGAAAC
701 GGATTGATGA ACTGATTGCC GACGGCACGC TGTATGGCGG TATGCTGCCG
751 AAAATCGCTT CTGCGGTCGA AGCCGCCGTC AACGGTGTGA AAGCCACGCA
    TATCATCGAC GGCAGGTTGC CCAACGCGCT TTTGCTGGAA ATCTTTACCG
    ATGCCGGTAT CGGTTCGATG ATTTTGGGCG GTGGGGAAGA TGCCTGA
```

This corresponds to the amino acid sequence <SEQ ID 538; ORF 138>:

1 MESENIISAA DKARILAEAL PYIRRFSGSV AVIKYGGNAM TEPALKEGFA

```
51 RDVVLLKLVG IHPVIVHGGG PQINAMLEKV GKKGEFVOGM RVTDKEAMDI
          VEMVLGGHVN KEIVSMINTY GGHAVGVSGR DDHFIKAKKL LIDTPEQNGV
          DIGQVGTVES IDTGLVKGLI ERGCIPVVAP VGVGEKGEAF NINADLVAGK
          LAEELNAEKL LMMTNIAGVM DKTGNLLTKL TPKRIDELIA DGTLYGGMLP
          KIASAVEAAV NGVKATHIID GRLPNALLLE IFTDAGIGSM ILGGGEDA*
Computer analysis of this amino acid sequence gave the following results:
Homology with a predicted ORF from N. gonorrhoeae
ORF 138 shows 98.0% identity over a 298 aa overlap with a predicted ORF (ORF 138.ng)
from N. gonorrhoeae:
m138/g138
                      10
                                20
                                           30
                                                               50
             MESENIISAADKARILAEALPYIRRFSGSVAVIKYGGNAMTEPALKEGFARDVVLLKLVG
m138.pep
                                20
                                           30
                                                     40
                                                               50
```

```
MEFENIISAADKARILAEALPYIRRFSGSVAVIKYGGNAMTEPALKEGFARDVVLLKLVG
g138
                70
                       80
                               90
                                      100
                                             110
                                                     120
m138.pep
          IHPVIVHGGGPQINAMLEKVGKKGEFVQGMRVTDKEAMDIVEMVLGGHVNKEIVSMINTY
          IHPVIVHGGGPQINAMLEKVGKKGEFVQGMRVTDKETMDIVEMVLGGHVNKEIVSMINTY
g138
                70
                       80
                               90
                                      100
                                             110
               130
                       140
                              150
                                      160
                                             170
                                                     180
m138.pep
          GGHAVGVSGRDDHFIKAKKLLIDTPEQNGVDIGQVGTVESIDTGLVKGLIERGCIPVVAP
          g138
          GGHAVGVSGRDDHFIKAKKLLVDTPEQNSVDIGQVGTVESIDTGLVKGLIERGCIPVVAP
               130
                       140
                              150
                                     160
                                             170
                                                     180
               190
                       200
                              210
                                     220
                                             230
                                                     240
m138.pep
         VGVGEKGEAFNINADLVAGKLAEELNAEKLLMMTNIAGVMDKTGNLLTKLTPKRIDELIA
          g138
          VGVGEKGEAFNINADLVAGKLAEELNAEKLLMMTNIAGVMDKTGNLLTKLTPKRIDGLIA
               190
                       200
                              210
                                     220
                                             230
               250
                      260
                              270
                                     280
                                             290
                                                    299
         DGTLYGGMLPKIASAVEAAVNGVKATHIIDGRLPNALLLEIFTDAGIGSMILGGGEDAX
m138.pep
          DGTLYGGMLPKIASAVEAAVNGVKATHIIDGRLPNALLLEIFTDAGIGSMILGRGEDAX
g138
               250
                       260
                              270
                                     280
                                             290
```

The following partial DNA sequence was identified in N. meningitidis <SEO ID 539>:

```
a138.seq
          ATGGAGTCTG AAAACATTAT TTCCGCCGCC GACAAGGCGC GTATCCTTGC
       1
          CGAAGCGCTG CCTTACATCC GCCGGTTTTC CGGTTCGGTC GCCGTCATCA
      51
          AATACGGCGG CAACGCGATG ACCGAACCTG CCTTGAAAGA AGGGTTTGCC
     101
          CGCGATGTCG TGCTGCTGAA GCTGGTCGGC ATTCATCCCG TCATCGTTCA
     201
          CGGCGGCGGG CCGCAGATCA ATGCGATGCT TGAAAAAGTC GGCAAAAAGG
          GTGAGTTTGT CCAAGGAATG CGCGTTACCG ACAAAGAGGC GATGGATATT
     251
     301
          GTCGAAATGG TGTTGGGCGG GCATGTCAAT AAAGAAATCG TGTCGATGAT
          TAACACATAT GGCGGACACG CGGTCGGCGT AAGCGGACGC GACGACCATT
     351
          TCATTAAGGC GAAGAAACTT TTGATCGATA CGCCCGAACA GAATGGCGTG
     401
     451
          GACATCGGAC AGGTCGGTAC GGTGGAAAGC ATCGATACCG GTTTGGTTAA
     501
          AGGGCTGATA GAACGTGGCT GCATTCCCGT CGTCGCCCCC GTCGGCGTAG
     551
          GTGAAAAAGG CGAAGCGTTC AACATCAACG CCGATTTGGT AGCAGGCAAA
          TTGGCGGAAG AATTGAACGC CGAAAAACTC TTGATGATGA CGAATATCGC
     601
          CGGTGTGATG GACAAAACGG GCAATCTGCT GACCAAACTC ACGCCGAAAC
     651
          GGATTGATGA ACTGATTGCC GACGGCACGC TGTATGGCGG TATGCTGCCG
     701
     751
          AAAATCGCTT CTGCGGTCGA AGCCGCCGTC AACGGCGTGA AAGCCACGCA
          TATCATCGAC GGCAGGGTGC CCAACGCGCT TTTGCTGGAA ATCTTTACCG
     801
     851
          ATGCCGGTAT CGGTTCGATG ATTTTGGGCG GTGGGGAAGA TGCCTGA
```

BNSDOCID: <WO 9957280A2 1 >

```
This corresponds to the amino acid sequence <SEQ ID 540; ORF 138.a>:
       a138.pep
                MESENIISAA DKARILAEAL PYIRRFSGSV AVIKYGGNAM TEPALKEGFA
                RDVVLLKLVG IHPVIVHGGG PQINAMLEKV GKKGEFVQGM RVTDKEAMDI
             51
                VEMVLGGHVN KEIVSMINTY GGHAVGVSGR DDHFIKAKKL LIDTPEQNGV
            101
                DIGOVGTVES IDTGLVKGLI ERGCIPVVAP VGVGEKGEAF NINADLVAGK
            151
                LAEELNAEKL LMMTNIAGVM DKTGNLLTKL TPKRIDELIA DGTLYGGMLP
            201
                KIASAVEAAV NGVKATHIID GRVPNALLLE IFTDAGIGSM ILGGGEDA*
            251
            99.7% identity in 298 aa overlap
  m138/a138
                          10
                                   20
                   MESENIISAADKARILAEALPYIRRFSGSVAVIKYGGNAMTEPALKEGFARDVVLLKLVG
                                                     40
                                                              50
      m138.pep
                   MESENIISAADKARILAEALPYIRRFSGSVAVIKYGGNAMTEPALKEGFARDVVLLKLVG
      a138
                                   20
                                            30
                                                              50
                                                                       60
                          70
                                   80
                                            90
                                                    100
      m138.pep
                  IHPVIVHGGGPQINAMLEKVGKKGEFVQGMRVTDKEAMDIVEMVLGGHVNKEIVSMINTY
                                                             110
                                                                      120
                  IHPVIVHGGGPQINAMLEKVGKKGEFVQGMRVTDKEAMDIVEMVLGGHVNKEIVSMINTY
      a138
                         70
                                  80
                                            90
                                                    100
                                                             110
                                                                      120
                        130
                                  140
                                          150
                                                    160
                  GGHAVGVSGRDDHFIKAKKLLIDTPEQNGVDIGQVGTVESIDTGLVKGLIERGCIPVVAP
                                                             170
      m138.pep
                  GGHAVGVSGRDDHFIKAKKLLIDTPEQNGVDIGQVGTVESIDTGLVKGLIERGCIPVVAP
      a138
                                 140
                                          150
                                                   160
                                                            170
                        190
                                 200
                                          210
                                                   220
                  VGVGEKGEAFNINADLVAGKLAEELNAEKLLMMTNIAGVMDKTGNLLTKLTPKRIDELIA
      m138.pep
                  a138
                  VGVGEKGEAFNINADLVAGKLAEELNAEKLLMMTNIAGVMDKTGNLLTKLTPKRIDELIA
                        190
                                 200
                                          210
                                                   220
                                                            230
                                                                     240
                        250
                                 260
                                          270
                                                   280
                                                            290
                 DGTLYGGMLPKIASAVEAAVNGVKATHIIDGRLPNALLLEIFTDAGIGSMILGGGEDAX
     m138.pep
                 DGTLYGGMLPKIASAVEAAVNGVKATHIIDGRVPNALLLEIFTDAGIGSMILGGGEDAX
     a138
                                 260
                                          270
                                                   280
                                                            290
The following partial DNA sequence was identified in N.gonorrhoeae <SEQ ID 541>:
     g139.seq
              ATGCGAACCA CCTCAACCTT CCCTACAAAA ACTTTCAAAC CGGCTGCCAT
            1
              GGCGTTAGCT GTTGCAACAA CACTTTCTGC CTGCTTAggc ggcggcggag
           51
              GCGGCACTTC TGCTCCCGAC TTTAATGCAG GCGGCACCGG TATCGGCAGC
          101
              AACAGCAGGG CAACGATAGC GGAATCAGCA GCAGTATCTT ACGCCGGTAT
          151
              AAAAAACGAA ATGTGCAAAG ACAGAAGCAT GCTCTGTGCC GGTCGGGATG
          201
              ACGTTGCGGT TACAGACAGG GATGCCAAAA TCAAAGCCCC CCGAATCTGC
         251
              ATACCGGAGA CTTTTCAAAC CCAAATGACC AATATTAAGA ATATGATCAA
         301
              CCTCAAACCT GCAATTGAAG CAGGCTATAC AGGACGCGGG GTAGAGGTAG
         351
              GTATCGTCGA TACAGGCGAA TCCGTCGGCA GCATATCCTT TCCCGAACTG
         401
              TATGGCAGAA AAGAACACGG CTATAACGAA AATTACAAAA ACAAATTACA
              AAAACTATAC GGCGTATATG CGGAAGGAAG CGCCTGA
         501
This corresponds to the amino acid sequence <SEQ ID 542; ORF 138.ng>:
    g139.pep
             MRTTSTFPTK TFKPAAMALA VATTLSACLG GGGGGTSAPD FNAGGTGIGS
           1
             NSRATIAESA AVSYAGIKNE MCKDRSMLCA GRDDVAVTDR DAKIKAPRIC
          51
```

IPETFQTQMT NIKNMINLKP AIEAGYTGRG VEVGIVDTGE SVGSISFPEL

151 YGRKEHGYNE NYKNKLQKLY GVYAEGSA*

101

```
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 543>:
     m139.seq
           1
              ATGCGAACGA CCCCAACCTT CCCTACAAAA ACTTTCAAAC CGACTGCCAT
          51 GGCGTTAGCT GTTGCAACAA CACTTTCTGC CTGCTTAGGC GGCGGCGGAG
          101 GCGGCACTTC TGCGCCCGAC TTCAATGCAG GCGGTACCGG TATCGGCAGC
          151 AACAGCAGAG CAACAACAGC GAAATCAGCA GCAGTATCTT ACGCCGGTAT
         201 CAAGAACGAA ATGTGCAAAG ACAGAAGCAT GCTCTGTGCC GGTCGGGATG
         251 ACGTTGCGGT TACAGACAGG GATGCCAAAA TCAATGCCCC CCCCGAATC
         301 TGCATACCGG AGACTTTCCA AACCCAAATG ACGCATLACA AGAATTTGAT
         351 CAACCTCAAA CCTGCAATTG AAGCAGGCTA TACAGGACGC GGGGTAGAGG
         401 TAGGTATCGT CGACACAGGC GAATCCGTCG GCAGCATATC CTTTCCCGAA
         451 CTGTATGGCA GAAAAGAACA CGGCTATAAC GAAAATTACG AAAAACTATA
         501 CGGCGTATAT GCGGAAGGAA GCGCCTGA
This corresponds to the amino acid sequence <SEQ ID 544; ORF 138>:
     m139.pep
              MRTTPTFPTK TFKPTAMALA VATTLSACLG GGGGGTSAPD FNAGGTGIGS
          51
              NSRATTAKSA AVSYAGIKNE MCKDRSMLCA GRDDVAVTDR DAKINAPPRI
              CIPETFQTQM THYKNLINLK PAIEAGYTGR GVEVGIVDTG ESVGSISFPE
          151 LYGRKEHGYN ENYEKLYGVY AEGSA*
Computer analysis of this amino acid sequence gave the following results:
Homology with a predicted ORF from N.gonorrhoeae
ORF 138 shows 92.2% identity over a 179 aa overlap with a predicted ORF (ORF 138.ng)
from N. gonorrhoeae:
     m139/g139
                         10
                                   20
                                                      40
                                                               50
                                                                         60
     m139.pep
                 MRTTPTFPTKTFKPTAMALAVATTLSACLGGGGGGTSAPDFNAGGTGIGSNSRATTAKSA
                 MRTTSTFPTKTFKPAAMALAVATTLSACLGGGGGGTSAPDFNAGGTGIGSNSRATIAESA
     g139
                                  20
                                            30
                                                      40
                                                               50
                         70
                                            90
                                                     100
                                                              110
     m139.pep
                 AVSYAGIKNEMCKDRSMLCAGRDDVAVTDRDAKINAPPRICIPETFQTQMTHYKNLINLK
                 AVSYAGIKNEMCKDRSMLCAGRDDVAVTDRDAKIKAP-RICIPETFQTQMTNIKNMINLK
     g139
                         70
                                  80
                                            90
                                                      100
                        130
                                  140
                                           150
                                                     160
                                                                  170
     m139.pep
                 PAIEAGYTGRGVEVGIVDTGESVGSISFPELYGRKEHGYNENY----EKLYGVYAEGSAX
                 : | | | | | | | | | | | |
     g139
                 PAIEAGYTGRGVEVGIVDTGESVGSISFPELYGRKEHGYNENYKNKLQKLYGVYAEGSAX
               120
                         130
                                  140
                                            150
                                                      160
                                                               170
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 545>:
     a139.seq
           1
              ATGCGAACGA CCCCAACCTT CCCTACAAAA ACTTTCAAAC CGGCTGCCAT
          51
              GGCGTTAGCT GTTGCAACAA CACTTTCTGC CTGCTTAGGC GGCGGCGGAG
         101 GCGGCACTTC TGCGCCCGAC TTCAATGCAG GCGGCACCGG TATCGGCAGC
              AACAGCAGGG CAACAACAGC GAAATCAGCA GCAATATCTT ACGCCGGTAT
         201
              CAAGAACGAA ATGTGCAAAG ACAGAAGCAT GCTCTGTGCC GGTCGGGATG
         251
              ACGTTGCGGT TACAGACAGG GATGCCAAAA TCAATGCCCC CCCCGAATC
         301
              TGCATACCGG AGACTTTACA AACCCAAATG ACGCAT.ACA AGAATTTGAT
             CAACCTCAAA CCTGCAATTG AAGCAGGCTA TACAGGACGC GGGGTAGAGG
         351
         401
              TAGGTATCGT CGACACAGGC GAATCCGTCG GCAGCATATC CTTTCCCGAA
         451
             CTGTATGGCA GAAAAGAACA CGGCTATAAC GAAAATTAC. AAAAACTATA
             CGGCGTATAT GCGGAAGGAA GCGCCTGA
```

This corresponds to the amino acid sequence <SEQ ID 546; ORF 139.a>: a139.pep

¹ MRTTPTFPTK TFKPAAMALA VATTLSACLG GGGGGTSAPD FNAGGTGIGS

- 51 NSRATTAKSA AISYAGIKNE MCKDRSMLCA GRDDVAVTDR DAKINAPPRI CIPETLOTOM THXKNLINLK PAIEAGYTGR GVEVGIVDTG ESVGSISFPE 101
- 151 LYGRKEHGYN ENYXKLYGVY AEGSA*

## m139/a139 97.1% identity in 175 aa overlap

m139.pep	10 MRTTPTFPTFPT	20	30	40	50	60
	MRTTPTFPTKTFKP	TAMALAVAT	TLSACLGGGGG	GTSAPDFNAG	GTGIGSNSR	ATTAKSA
a139	MRTTPTFPTKTFKP		1   4   1   1   1   1   1			
	10	20	30	GISAPDENAG		ATTAKSA
	- •	20	30	40	50	60
	70	80	90	100	110	
m139.pep	AVSYAGIKNEMCKDI	RSMLCAGRDE	איני מח פחידעו בעו	NADDDICTOR	110	120
	1:111111111111					CNLINLK
<b>a</b> 139					1:11111	
	AISYAGIKNEMCKDI 70	CAGRDL	VAVTDRDAKI	NAPPRICIPE	<b>FLOTOMTHX</b> K	NLINLK
	70	80	90	100	110	120
100	130	140	150	160	170	
m139.pep	PAIEAGYTGRGVEVO	SIVDTGESVG	SISFPELYGR	KEHCVNENVE	T VCIUNDOS	
	111111111111111	11111111		KENGINENIEI	LIGVYAEGS	AX
a139	PAIEAGYTGRGVEVO	TVDTCESVC	CTCEDETWON	1111111	1111111111	11
	PAIEAGYTGRGVEVO	140	SISPPELYGR]	KEHGYNENYXI	<b>KLYGVYAEGS</b>	AX
	130	140	150	160	170	

#### The following partial DNA sequence was identified in N.gonorrhoeae <SEQ ID 547>: g140.seq

```
Atgtcggcac gCGGCAAGGG GGCAGgctat ctcAACAGTA CCGGACGACa
    1
      TGTTCCCTTC CTGAGTGCCG CCAAAATCGG GCAGGATTAT TCTTTCTTCA
 101 AAAATATCAA AACCGACGGC GGTCTGCTGG CTTCCCTCGA CAGCGTCGAA
 151 AAAACAGCGG GCAGTGAAGG CGACACGCCG TCCTATTATG TCCGTCGCGG
 201 CAATGCGGCA CGGACTGCTT CGGCAGCGGC ACATTCCGCG CCCGCCGGTC
      TGAAACACGC CGTAGAACAG GGCGGCAGCA ATCTGGAAAA CCTGATGGTC
      GAGCTGGATG CCTCCGAATC ATCCGCAACA CCCGAGACGG TTGAAACTGC
      GGTCGCCGAC CGCACAGATA TGCCGGGCAT CCGCCTACGG CGCACAACTT
 351
 401 TCCGCACAGC GGCAGCCGTA CAGCATGCGA ATACCGCCGA CGGCGTACGC
 451 aTCTTcaaCA GTCTCGCCGC TAccgTCTAt GccgACAGTG CCGCCGCCCA
 501 TGCCGATATG CAGGGACGCC GCCTGAAAGC CGTATCGGAC GGGTTGGACC
 551 ACAACGGTAC GGGTCTGCGC GTCATCGCGC AAACCCAACA GGACGGTGGA
 601 ACGTGGGAAC AGGGCGGTGT CGAAGGCAAA ATGCGCGGCA GTACCCAAAC
 651 TATCGGCATT GCCGCGAAAA CCGGCGAAAA TACGACAGCA GCCGCCACAC
 701 TGGGCATAGG ACGCAGCACA TGGAGCGAAA ACAGTGCAAA TGCAAAAACC
 751 GACAGCATTA GTCTGTTTGC AGGCATACGG CACGATGTGG GCGATATCGG
     CTATCTCAAA GGCCTGTTCT CctaCGGACG CTACAAAAAC AGCATCAGCC
 851 GCAGCACCGG TGCGGATGAA TATGCGGAAG GCAGCGTCAA CGGCACGCTG
 901 ATGCAGCTGG GCGCACTGGG TGGTGTCAAC GTTCCGTTTG CCGCAACGGG
 951 AGATTTGACG GTTGAAGGCG GTCTGCGCCA CGACCTGCTC AAACAGGATG
1001 CATTCGCCGA AAAAGGCAGT GCTTTGGGCT GGAGCGGCAA CAGCCTCACT
1051 GAAGGCACAC TGGTCGGACT CGCGGGTCTG AAACTGTCGC AACCCTTGAG
1101 CGATAAAGCC GTCCTGTCTG CGACGGCGGG CGTGGAACGC GACCTGAACG
1151 GACGCGACTA CGCGGTAACG GGCGGCTTTA CCGGCGCGGC TGCAGCAACC
1201 GGCAAGACGG GTGCACGCAA TATGCCGCAC ACCCGCCGGG TTGCCGGTCT
     GGGGGTGGAT GTCGAATTCG GCAACGGCTG GAACGGCTTG GCACGTTACA
1251
     GCTACACCGG TTCCAAACAG TACGGCAACC ACAGCGGACA AATCGGCGTA
1351
     GGCTACCGGT TCTGA
```

#### This corresponds to the amino acid sequence <SEQ ID 548; ORF 140.ng>: g140.pep

1	MSARGKGAGY	LNSTGRHVPF	LSAAKIGODY	SEPVALL VODO	GT T 7 GT 7 GT
51	KTAGSEGDTP	SYYVRRGNAA	RTASAAAHSA	DACT MINITED	GLLASLDSVE
101	ELDASESSAT	PETVETAVAD	PTDMDCIDID	PAGLICHAVEO	GGSNLENLMV
151	IFNSLAATVY	ADSAAAHADM	OCERT VALUE	RITTRIAAAV	QHANTADGVR
201	TWEOGGVEGK	MRGSTOTTCT	DGKKLKAVSD	GLDHNGTGLR	VIAQTQQDGG
	TWEQGGVEGK	.4600101161	AAKTGENTTA	AATLGIGRST	WSENSANAKT

251 DSISLFAGIR HDVGDIGYLK GLFSYGRYKN SISRSTGADE YAEGSVNGTL

```
MQLGALGGVN VPFAATGDLT VEGGLRHDLL KQDAFAEKGS ALGWSGNSLT
               EGTLVGLAGL KLSQPLSDKA VLSATAGVER DLNGRDYAVT GGFTGAAAAT
               GKTGARNMPH TRRVAGLGVD VEFGNGWNGL ARYSYTGSKQ YGNHSGQIGV
          451
               GYRF*
The following partial DNA sequence was identified in N. meningitidis <SEO ID 549>:
     m140.seq
              ATGTCGGCAC GCGGCAAGGG GGCAGGCTAT CTCAACAGTA CCGGACGACG
            1
           51
              TGTTCCCTTC CTGAGTGCCG CCAAAATCGG GCAGGATTAT TCTTTCTTCA
          101
              CAAACATCGA AACCGACGGC GGCCTGCTGG CTTCCCTCGA CAGCGTCGAA
              AAAACAGCGG GCAGTGAAGG CGACACGCTG TCCTATTATG TCCGTCGCGG
               CAATGCGGCA CGGACTGCTT CGGCAGCGGC ACATTCCGCG CCCGCCGGTC
               TGAAACACGC CGTAGAACAG GGCGGCAGCA ATCTGGAAAA CCTGATGGTC
          301
               GAACTGGATG CCTCCGAATC ATCCGCAACA CCCGAGACGG TTGAAACTGC
          351
               GGCAGCCGAC CGCACAGATA TGCCGGGCAT CCGCCCCTAC GGCGCAACTT
          401 TCCGCGCAGC GGCAGCCGTA CAGCATGCGA ATGCCGCCGA CGGTGTACGC
          451 ATCTTCAACA GTCTCGCCGC TACCGTCTAT GCCGACAGTA CCGCCGCCCA
          501 TGCCGATATG CAGGGACGCC GCCTGAAAGC CGTATCGGAC GGGTTGGACC
          551 ACAACGGCAC GGGTCTGCGC GTCATCGCGC AAACCCAACA GGACGGTGGA
          601 ACGTGGGAAC AGGGCGGTGT TGAAGGCAAA ATGCGCGGCA GTACCCAAAC
          651 CGTCGGCATT GCCGCGAAAA CCGGCGAAAA TACGACAGCA GCCGCCACAC
          701 TGGGCATGGG ACGCAGCACA TGGAGCGAAA ACAGTGCAAAA TGCAAAAACC
          751 GACAGCATTA GTCTGTTTGC AGGCATACGG CACGATGCGG GCGATATCGG
          801 CTATCTCAAA GGCCTGTTCT CCTACGGACG CTACAAAAAC AGCATCAGCC
          851 GCAGCACCGG TGCGGACGAA CATGCGGAAG GCAGCGTCAA CGGCACGCTG
          901 ATGCAGCTGG GCGCACTGGG CGGTGTCAAC GTTCCGTTTG CCGCAACGGG
          951 AGATTTGACG GTCGAAGGCG GTCTGCGCTA CGACCTGCTC AAACAGGATG
         1001 CATTCGCCGA AAAAGGCAGT GCTTTGGGCT GGAGCGGCAA CAGCCTCACT
         1051 GAAGGCACGC TGGTCGGACT CGCGGGTCTG AAGCTGTCGC AACCCTTGAG
         1101 CGATAAAGCC GTCCTGTTTG CAACGGCGGG CGTGGAACGC GACCTGAACG
         1151 GACGCGACTA CACGGTAACG GGCGGCTTTA CCGGCGCGAC TGCAGCAACC
         1201 GGCAAGACGG GGGCACGCAA TATGCCGCAC ACCCGTCTGG TTGCCGGCCT
         1251 GGGCGCGGAT GTCGAATTCG GCAACGGCTG GAACGGCTTG GCACGTTACA
               GCTACGCCGG TTCCAAACAG TACGGCAACC ACAGCGGACG AGTCGGCGTA
               GGCTACCGGT TCTGA
This corresponds to the amino acid sequence <SEQ ID 550; ORF 140>:
     m140.pep
            1 MSARGKGAGY LNSTGRRVPF LSAAKIGQDY SFFTNIETDG GLLASLDSVE
           51 KTAGSEGDTL SYYVRRGNAA RTASAAAHSA PAGLKHAVEQ GGSNLENLMV
          101 ELDASESSAT PETVETAAAD RTDMPGIRPY GATFRAAAAV QHANAADGVR
          151 IFNSLAATVY ADSTAAHADM QGRRLKAVSD GLDHNGTGLR VIAQTQQDGG
          201 TWEQGGVEGK MRGSTQTVGI AAKTGENTTA AATLGMGRST WSENSANAKT
          251 DSISLFAGIR HDAGDIGYLK GLFSYGRYKN SISRSTGADE HAEGSVNGTL
          301 MQLGALGGVN VPFAATGDLT VEGGLRYDLL KODAFAEKGS ALGWSGNSLT
          351 EGTLVGLAGL KLSQPLSDKA VLFATAGVER DLNGRDYTVT GGFTGATAAT
              GKTGARNMPH TRLVAGLGAD VEFGNGWNGL ARYSYAGSKQ YGNHSGRVGV
               GYRF*
Computer analysis of this amino acid sequence gave the following results:
Homology with a predicted ORF from N.gonorrhoeae
ORF 140 shows 94.5% identity over a 454 as overlap with a predicted ORF (ORF 140.ng)
from N. gonorrhoeae:
     m140/g140
                                    20
                                              30
                                                       40
                  {\tt MSARGKGAGYLNSTGRRVPFLSAAKIGQDYSFFTNIETDGGLLASLDSVEKTAGSEGDTL}
     m140.pep
                  MSARGKGAGYLNSTGRHVPFLSAAKIGQDYSFFKNIKTDGGLLASLDSVEKTAGSEGDTP
     g140
                                             30
                                                       40
                                                                 50
                                                                           60
                          70
                                    80
                                             90
                                                      100
                                                                110
                                                                          120
```

WO 99/57280



m140.pep	SYYVRRGNAARTA	SAAAHSAPAC	SLKHAVEQGGS	NLENLMVELD	ASESSATDE	TT I DOWN A A S.
g140						
	70	80	90	100	ASESSATPE 110	IVETAVAD 120
m140.pep	130 RTDMPGIRPYCar	140	150	160	170	180
	RTDMPGIRPYGAT	H: HIHIH	NAADGVRIFN:	SLAATVYADS:	TAAHADMOGI	RLKAVSD
<b>g14</b> 0	RTDMPGIRLRRTT	FRTAAAVOHA	:		:	
	130	140	150	160	жанармодг 170	
				100	170	180
m140	190	200	210	220	230	240
m140.pep	GLDHNGTGLRVIAC	)TQQDGGTWE	QGGVEGKMRGS	STOTVGIAAKT	GENTTAAAT	T-CACDAD-T
g140						
3-10	GLDHNGTGLRVIAC	)TQQDGGTWE(	DGGA F.GKWKGS	TOTIGIAAKT	GENTTAAAT	LGIGRST
	190	200	210	220	230	240
	250	260	270			
m140.pep	WSENSANAKTDSIS	LFAGIRHDAC	27U EDIGVI.KGI RO	280	290	300
g140	WSENSANAKTDSIS	LFAGIRHDVC	DIGYLKGLES	YGRYKNSTED		111111
	250	260	270	280	290	300
					250	300
m140.pep	310	320	330	340	350	360
штто.рер	MQLGALGGVNVPFA	ATGDLTVEGG	LRYDLLKQDA	FAEKGSALGW:	SGNSLTEGTI	VGLAGL
g140						
_	MQLGALGGVNVPFA	320	LRHDLLKQDAI	FAEKGSALGW		VGLAGL
		320	330	340	350	360
	370	380	390	400	43.0	
m140.pep	KLSQPLSDKAVLFA	AGVERDLNG	RDYTVTGGFT	ATAATGKTG	410	420
~1.40						
g140	KLSOPLSDKAVLSAT	SAUTION STATES	RDYAVTGGFTG	AAAATGKTGA	RNMPHTRRV	AGLGVD
	370	380	390	400	410	420
	430	440				
m140.pep	VEFGNGWNGLARYSY	440	450			
		:	>GKVGVGYRFX			
g140	VEFGNGWNGLARYSY	TGSKOYGNI	[			
	430	440	450			

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 551>:

	AGATTTGACG GTCGAAGGCCCATTCGCCGA AAAAGGCACGCGATAAAGCC GTCCTGTTTCGACGCCGAAGACGCGCAAGCAGCGCAAGCAGCGCAAGCAGC	GCTTTGGCT CGCGGGTCTG CAACGGCGGG GGCGGCTTTA TATGCCGCAC GGCAACGGCTG TACGGCAACC	GGAGCGGCAA AAGCTGTCGC CGTGGAACGC CCGGCGCGAC ACCCGCCTGG GAACGGCTTG ACAGCGGACG	CAGCATCACT AACCCTTGAG GACCTGAACG TGCAGCAACC TTGCCGGTCT GCACGTTACA AGTCGGCGTA	
a140.pep	MCACCECACY INDECODUR				
1 51 101 151 201	MSAGGKGAGY LNRTGQRVPI KTAGSEGDTL SYYVRRGNAI ELDASESSAT PETVETAAAI IFNNLAATVY ADSTAAHADI TWEQGGVEGK MRGSTQTVG	A RTASAAAHSA D RTDMPGIRPY M QGRRLKAVSD I AAKTGENTTA	PAGLKHAVEQ GATFRAAAAV GLDHNATGLR AATLGMGHST	GGSNLENLMV QHANAADGVR VIAQTQQDGG WSENSANAKT	
251	DSISLFAGIR HDAGDIGYLE	GLFSYGRYKN	SISRSTGADE	HAEGSVNGTL	
301 351	MQLGALGGVN VPFAATGDLT EGTLVGLAGL KLSQPLSDK	VEGGLRYDLL	KQDAFAEKGS	ALGWSGNSIT	
401 451	GKTGARNMPH TRLVAGLGAI GYRF*	VEFGNGWNGL	ARYSYAGSKQ	YGNHSGRVGV	
m140/a140 98	3.2% identity in 454 aa o	verlap	0 40	50	60
m140.pep	MSARGKGAGYLNSTGRRY				60 EGDTL
a140	:   MSAGGKGAGYLNRTGQRV 10	111111111111111111111111111111111111111	 YSFFTNIETDGG	1111111111111111	11111
					•
m140.pep	70 SYYVRRGNAARTASAAA 	80 9( SAPAGLKHAVEÇ	OGGSNLENLMVE	110 LDASESSATPETVE	120 TAAAD
a140	SYYVRRGNAARTASAAAI	ISAPAGLKHAVE(	OGGSNLENLMVE	LDASESSATPETVE	TAAAD
	70	80 90	100	110	120
m140.pep	130 RTDMPGIRPYGATFRAA	.40 150 AVQHANAADGVI		170 ADSTAAHADMOGRRL	180 KAVSD
	11111111111111111	111111111111		1111111111111111	11111
a140	RTDMPGIRPYGATFRAAA 130 1	AAVQHANAADGVI .40 150		ADSTAAHADMQGRRI 170	KAVSD 180
		200 210		230	240
m140.pep	GLDHNGTGLRVIAQTQQI	GGTWEQGGVEG	MRGSTOTVGIA	AKTGENTTAAATLG	MGRST
a140	:				
		210	-	230	240
m140.pep	WSENSANAKTDSISLFAC	60 270 SIRHDAGDIGYLE	KGLFSYGRYKNS	290 ISRSTGADEHAEGS	300 VNGTL
-140		11111111111			11111
a140	WSENSANAKTDSISLFAG 250 2	:60 270		ISRSTGADEHAEGS 290	300
		20 330		350	360
m140.pep	MQLGALGGVNVPFAATGI	LTVEGGLRYDLI	LKQDAFAEKGSA	LGWSGNSLTEGTLV	GLAGL
a140	MQLGALGGVNVPFAATGI	LTVEGGLRYDLI	LKODAFAEKGSA	!!!!!!!:!!!!! LGWSGNSITEGTLV	  GLAGL
		20 330		350	360
m140.pep	370 KLSQPLSDKAVLFATAGV	80 390 ERDLNGRDYTVI		410 KTGARNMPHTRI.VA	420 GLGAD
		111111111111	11111111111	11111111111111	TITLE
a140	KLSQPLSDKAVLFATAGV	ERDLNGRDYTVI	GGFTGATAATG	KTGARNMPHTRLVA	GLGAD

BNSDOCID: <WO___9957280A2_j_>

398

370 380 390 400 410 420 430 440 450 m140.pep VEFGNGWNGLARYSYAGSKQYGNHSGRVGVGYRFX a140 VEFGNGWNGLARYSYAGSKQYGNHSGRVGVGYRFX 430 440

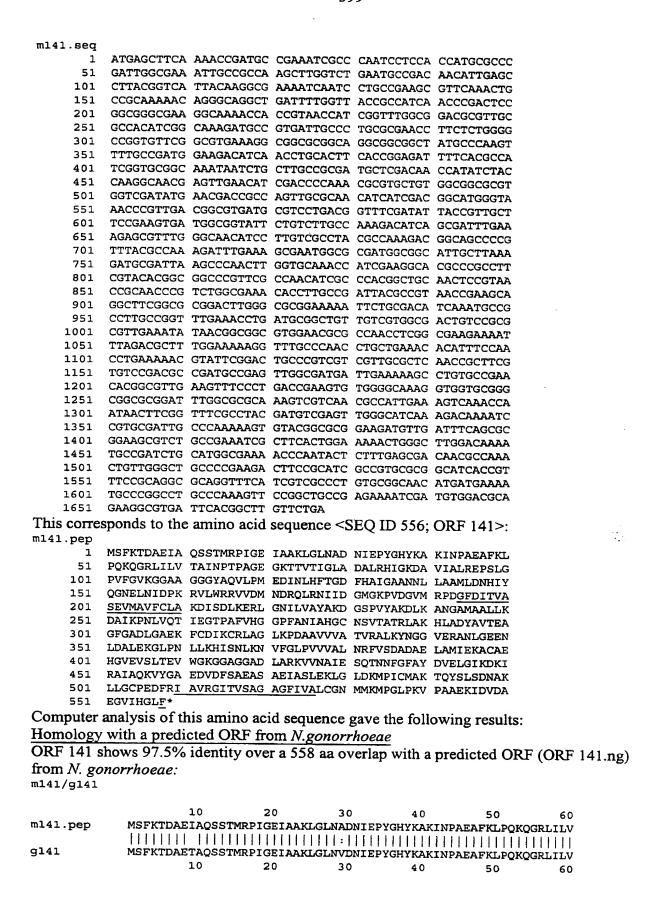
The following partial DNA sequence was identified in N.gonorrhoeae <SEQ ID 553>:

```
atgagettea aAAcegATGC CGAAACCGCC CAATCCTCCA CCATGCGCCC
    1
       GATTGGCGAA ATTGCCGCCA AGCTGGGTTT GAACGTTGAC AACATTGAGC
   51
       CTTACGGTCA TTACAAAGCC AAAATCAATC CTGCCGAAGC GTTCAAGCTG
  101
       CCGCAAAAAC AAGGCAGGCT GATTTTGGTT ACCGCCATCA ACCCGACTCC
      GGCGGGCGAA GGCAAAACCA CCGTAACCAT CGGTTTGGCG GACGCATTGC
  201
      GCCATATCGG CAAAGACTCT GTGATTGCTT TGCGCGAGCC TTCTTTGGGT
       CCGGTGTTCG GCGTGAAAGG CGGCGCGCA GGCGGCGGCT ACGCGCAAGT
      TTTGCCGATG GAAGACATCA ACCTGCACTT CACCGGCGAC TTCCACGCCA
      TCGGTGCGGC GAATAACCTC CTCGCCGCCA TGCTCGACAA CCATATCTAC
 401
      CAAGGTAACG AGTTGAACAT CGACCCCAAA CGCGTGCTGT GGCGGCGCGT
 451
      GGTCGATATG AACGACCGCC AGTTGCGCAA CATCATCGAC GGTATGGGCA
      AGCCTGTtga cggCGTGATG CGtcccGACG GCTTCGACAT CACCGTCGCC
      TCCGAAGTGa tggcgGTATT CTGCCTTGCC AAAGACATCA GCGATTTGAA
 601
      AGAGCGTTLL GGCAATATTC TCGTCGCCTA CGCCAAAGAC GGCAGCCCCG
 651
      TTTACGCCAA AGATTTGAAG GCACACGGCG CGATGGCGGC ATTGCTAAAA
 701
      GATGCGATTA AGCCCAATTT GGTGCAAACC ATCGAAGGCA CTCCGGCCTT
      TGTACACGGC GGCCCGTTCG CCAACATCGC CCACGGCTGC AACTCCGTTA
 801
      CCGCAACCCG TCTGGCGAAA CACCTTGCCG ATTACGCCGT AACCGAAGCA
 851
      GGCTTCGGCG CGGACTTGGG TGCGGAAAAA TTCTGCGACA TCAAATGCCG
 901
      CCTTGCCGGT TTGAAACCTG ATGCGGCAGT CGTCGTGGCG ACTGTCCGCG
 951
      CCCTGAAATA CAACGGCGGC GTGGAACGCG CCAACCTTGG TGAAGAAAAC
      CTCGAAGCCT TGGCAAAAGG TTTGCCCAAC CTGTTGAAAC ACATTTCCAA
1051
     CCTGAAAAAC GTATTCGGAC TGCCCGTCGT CGTTGCGCTC AACCGCTTCG
1101
     TGTCCGACTC CGATGCCGAG TTGGCGATGA TTGAAAAAGC CTGTGCCGAA
1151
1201 CACGGCGTTG AAGTTTCCCT GACCGAAGTG TGGGGCAAAG GCGGCGCGGG
1251 CGGCGCGGAT TTGGCGCGCA AAGTCGTCAA TGCCATCGAC AACCAACCTA
1301 ATAACTTCGG TTTCGCCTAC GATGTCGAGT TGGGCATCAA AGACAAAATC
1351 CGTGCGATTG CCCAAAAAGT GTACGGCGCG GAAGATGTCG ATTTCAGCGC
1401 GGAAGCGTCT GCCGAAATCG CCTCGCTGGA AAAACTGGGC TTGGACAAAA
     TGCCGATCTG CATGGCGAAA ACCCAATATT CATTGAGCGA CAACGCCAAA
     CTCTTGGGCT GCCCGAAGG CTTCCGCATC GCCGTACGCG GTATCACTGT
     TTCCGCCGGC GCGGGCTTCA TCGTTGCGTT GTGCGGCAAT ATGATGAAAA
     TGCCGGGCCT GCCGAAAGTT CCGGCTGCCG AGAAAATCGA TGTGGACGAA
     CACGGCGTGA TTCACGGCTT GTTCTGA
```

This corresponds to the amino acid sequence <SEQ ID 554; ORF 141.ng>:

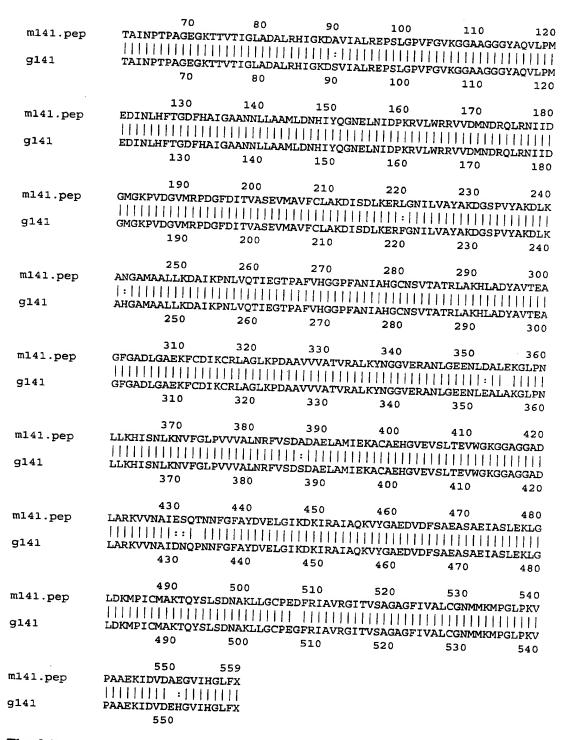
```
1 MSFKTDAETA QSSTMRPIGE IAAKLGLNVD NIEPYGHYKA KINPAEAFKL
51 PQKQGRLILV TAINPTPAGE GKTTVTIGLA DALRHIGKDS VIALREPSLG
101 PVFGVKGGAA GGGYAQVLPM EDINLHFTGD FHAIGAANNL LAAMLDNHIY
151 QGNELNIDPK RVLWRRVVDM NDRQLRNIID GMGKPVDGVM RPDGFDITVA
201 SEVMAVFCLA KDISDLKERF GNILVAYAKD GSPVYAKDLK AHGAMAALLK
251 DAIKPNLVQT IEGTPAFVHG GPFANIAHGC NSVTATRLAK HLADYAVTEA
301 GFGADLGAEK FCDIKCRLAG LKPDAAVVVA TVRALKYNGG VERANLGEEN
351 LEALAKGLPN LLKHISNLKN VFGLPVVVAL NRFVSDSDAE LAMIEKACAE
401 HGVEVSLTEV WGKGGAGGAD LARKVVNAID NQPNNFGFAY DVELGIKDKI
451 RAIAQKVYGA EDVDFSAEAS AEIASLEKLG LDKMPICMAK TQYSLSDNAK
501 LLGCPEGFRI AVRGITVSAG AGFIVALCGN MMKMPGLPKV PAAEKIDVDE
```

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 555>:



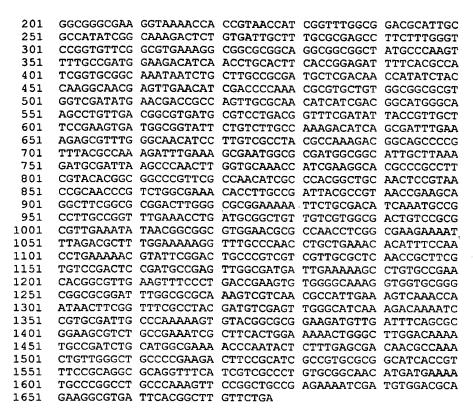
BNSDOCID: <WO___9957280A2_I_>

1.0



The following partial DNA sequence was identified in N. meningitidis <SEQ ID 557>:

101	CTTACGGTCA	TTACAAAGCC	AGCTGGGTTT	GAACGTTGAC	CCATGCGCCC	
	CITACGGICA	TTACAAAGCC	ΔΔΔΔΦΟλλπο	CMCCCCCTTCC	GTTCAAACTG ACCCGACTCC	



#### This corresponds to the amino acid sequence <SEQ ID 558; ORF 141.a>:

```
al41.pep
         MSFKTDAEIA OSSTMRPIGE IAAKLGLNVD NIEPYGHYKA KINPAEAFKL
      51
         PQKQGRLILV TAINPTPAGE GKTTVTIGLA DALRHIGKDS VIALREPSLG
         PVFGVKGGAA GGGYAQVLPM EDINLHFTGD FHAIGAANNL LAAMLDNHIY
    1.01
          QGNELNIDPK RVLWRRVVDM NDRQLRNIID GMGKPVDGVM RPDGFDITVA
    201
         SEVMAVFCLA KDISDLKERL GNILVAYAKD GSPVYAKDLK ANGAMAALLK
         DAIKPNLVQT IEGTPAFVHG GPFANIAHGC NSVTATRLAK HLADYAVTEA
    301
         GFGADLGAEK FCDIKCRLAG LKPDAAVVVA TVRALKYNGG VERANLGEEN
    351
         LDALEKGLPN LLKHISNLKN VFGLPVVVAL NRFVSDSDAE LAMIEKACAE
         HGVEVSLTEV WGKGGAGGAD LARKVVNAIE SQTNNFGFAY DVELGIKDKI
    451
         RAIAQKVYGA EDVDFSAEAS AEIASLEKLG LDKMPICMAK TQYSLSDNAK
         LLGCPEDFRI AVRGITVSAG AGFIVALCGN MMKMPGLPKV PAAEKIDVDA
         EGVIHGLF*
```

#### m141/a141 99.5% identity in 558 aa overlap

	10	20	30	40	50	60
m141.pep	MSFKTDAEIAQSST	MRPIGEIAAR	CLGLNADNIEP	YGHYKAKINP	AEAFKLPQKQ	GRLILV
		111111111	1111:1111	11111111111	11111111111	11111
a141	MSFKTDAEIAQSST	MRPIGEIAAK	LGLNVDNIEP	YGHYKAKINP	AEAFKLPQKQ	GRLILV
	10	20	30	40	50	60
	70	80	90	100	110	120
m141.pep	TAINPTPAGEGKTT	VTIGLADALE	RHIGKDAVIAL	REPSLGPVFG	VKGGAAGGGY	AQVLPM
			1111:111		1111111111	11111
a141	TAINPTPAGEGKTT	VTIGLADALF	RHIGKDSVIAL	REPSLGPVFG	VKGGAAGGGY	AQVLPM
	70	80	90	100	110	120
	130	140	150	160	170	180
m141.pep	EDINLHFTGDFHAI	GAANNLLAAN	ILDNHIYQGNE	LNIDPKRVLW	RRVVDMNDRC	LRNIID
		11111111	1111111111	1111111111	1111111111	11111
a141	EDINLHFTGDFHAI	GAANNLLAAM	LDNHIYQGNE	LNIDPKRVLW	RRVVDMNDRQ	LRNIID
	130	140	150	160	170	180

BNSDOCID: <WO__9957280A2_I_>

15 ° 5 1.78

WO 99/57280

	190	200	210			
m141.pep	GMGKPVDGVMRPDG	FDITVASEV	MAVECTARDT	220	230	240
a141		111111111		1111111		
a141		L DI I AWOFA	MAVFCLAKDI:	SDLKERLGNII	VAYAKDGSP	TIIIIII VYAKDIK
	190	200	210	220	230	240
m141.pep	250	260	270	280	290	300
mr41.pep	ANGAMAALLKDAIK	PNLVOTIEG	TPAFVHGGPF	ANIAHGCNSVI	ATRLAKHLA	
a141						
	ANGAMAALLKDAIK 250	260	270	NIAHGCNSVT 280	'ATRLAKHLAI	DYAVTEA
		200	270	280	290	<b>30</b> 0
m141.pep	310	320	330	340	350	360
pcp	GFGADLGAEKFCDI	KCRLAGLKP	DAAVVVATVRA	LKYNGGVERA	NLGEENLDAI	LEKGLPN
a141	GFGADLGAEKFCDTI			11111111	111111111	111111
	GFGADLGAEKFCDII 310	320	330	LKYNGGVERA 340	NLGEENLDAI	
			330	340	350	360
m141.pep	370	380	390	400	410	420
mr4r.pep	LLKHISNLKNVFGL	PVVVALNRF	SDADAELAMI:	EKACAEHGVE	TOT MELTING	
a141			11 *   1   1   1   1	111111111		
	LLKHISNLKNVFGLE 370	380	SUSDAELAMII 390	EKACAEHGVE		GAGGAD
		500	390	400	410	420
-141	430	440	450	460	470	480
m141.pep	LARKVVNAIESQTNN	FGFAYDVEL	GIKDKIRAIA	QKVYGAEDVD		SLEKLG
a141						
-	LARKVVNAIESQTNN 430	440	GIKDKIRAIA( 450	2KVYGAEDVDE	SAEASAEIA	SLEKLG
		440	450	460	470	480
2 4 2	490	500	510	520	530	E 4.0
m141.pep	LDKMPICMAKTQYSL	SDNAKLLGC	PEDFRIAVRGI	TUSAGAGETU	AT COMME	540 PGL PKV
a141		1111111	, , , , , , , , , , ,	1111111111		
	LDKMPICMAKTQYSL 490	SUNAKLLGC:	PEDFRIAVRGI	TVSAGAGFIV	ALCGNMMKMI	PGLPKV
	• • • • • • • • • • • • • • • • • • • •	300	510	520	530	540
-141	550	559				
m141.pep	PAAEKIDVDAEGVIH	GLFX				
a141	PARENTONAROUTE	1111				
	PAAEKIDVDAEGVIHO 550	<b>JLFX</b>				
	550					

#### The following partial DNA sequence was identified in N.gonorrhoeae <SEQ ID 559>: g142.seq

1	ATGCGTGCCG	ATTTCATGTT	TGCCGACAAT	ATCCCCCTCC	AGGTGCGCCA
51	ACGCGCC11C	TATTTCAAGT	TGTCCCGTTT	TCCCCCCATC	CCAAAMAMAA
101	TAGGCAAACC	GCTCTTCGGG	CGACAGGCCG	GTCAGCCCCC	CAAAAAA
151	GGCAACATCC	TGATGTTCGT	CCGCCAGCAT	ATTCATCCAC	7C~CMCCCC
201	TITCCGACAG	GAICGGAATG	AttcaCGCAC	Treatment A m	CCACACCARC
251	ACGGI CGGCG	GCTCGTCGGT	AACCGGCGCA	ACCGCCCTCA	TTCTT > TTCCC
301	GIMACGCCCT	GCCGCACCGT	CTGTCGTGAT	GACATGAACG	CCTCCCCCC
351	AGGATGCCAT	CGCATCACGG	AACGAAGTTT	CAAAACTUTUTUTUTUTUTUTUTUTUTUTUTUTUTUTUTU	CTCCAAATCC
401	GCCATTTTTC	CCCTTTAAAC	CGTCCCCTAT	ATAAGAATGC	TGCACACAAC
451	GCATCCCCC	ATGTGCAGCA	GTTCTGA		1 CONCINCAMO

This corresponds to the amino acid sequence <SEQ ID 560; ORF 142.ng>: g142.pep

- 1 MRADFMFADN MPVQVRQRAF YFKLSRFAAM PNMVGKPLFG RQAGQPGKMF
- 51 GNILMFVRQH IDAEAAVFRQ DRNDSRTPVY AQHHGRRLVG NRRNRRHCNA 101 VTPCRTVCRD DMNACRTGCH RITERSLKSF LQIRHFSPLN RPLYKNAAHK 151 ASPHVQQF*

The following partial DNA sequence was identified in N.meningitidis <SEQ ID 561>: m142.seq ATGCGTGCCG ATTTCATGTT TGCCGACAAT ATGCCCGTGC AGGTGCGCCA 1 ACGCGCCCTC TATTTCAAGT TGTCCCGTTT TGCCGCGATG CCAGATGTGG 51 TAGGCAAACC GCTCTTCGGG CGACAGGCCG GTCAGCCCGG CAAAATGTTC 101 GGCAACATCC TGATGTTCGT CCGCCAGCGT ATTGATGCAG AGGCTGCCGT 201 TTTCCGACAG GATCGGAATG ATTCGCGCAC TCCGGTTGAT GCACAGCATC 251 ACGGTCGGCG GCTCGTCGGT AACCGGCGGCG ACCGCCGTCA TTGTAATGCC 301 GTAACGCCCT GCCGCACCGT CTGTCGTGAT GACATGAACG CCTGCCGCGC 351 AAGATGCCAT CGCATCACGG AACGAAGTTT GAAAATTTTT CTGCAAATCC 401 GCCATTTTC CCCTTTAAAC TGTCCCCTAT ATAAGAATGC TGCACACAAG GCATCCCCC ATGTGCAGCA GTTTTGA This corresponds to the amino acid sequence <SEQ ID 562; ORF 142>: m142.pep MRADFMFADN MPVQVRQRAL YFKLSRFAAM PDVVGKPLFG RQAGQPGKMF GNILMFVRQR IDAEAAVFRQ DRNDSRTPVD AQHHGRRLVG NRRDRRHCNA 101 VTPCRTVCRD DMNACRARCH RITERSLKIF LQIRHFSPLN CPLYKNAAHK 151 ASPHVQQF* Computer analysis of this amino acid sequence gave the following results: Homology with a predicted ORF from N.gonorrhoeae ORF 142 shows 93.7% identity over a 158 aa overlap with a predicted ORF (ORF 142.ng) from N. gonorrhoeae: m142/g142 10 20 30 40 50 m142.pep MRADFMFADNMPVQVRQRALYFKLSRFAAMPDVVGKPLFGRQAGQPGKMFGN1LMFVRQR g142 MRADFMFADNMPVQVRQRAFYFKLSRFAAMPNMVGKPLFGRQAGQPGKMFGNILMFVRQH 10 20 30 40 50 80 90 100  ${\tt IDAEAAVFRQDRNDSRTPVDAQHHGRRLVGNRRDRRHCNAVTPCRTVCRDDMNACRARCH}$ m142.pep g142 IDAEAAVFRQDRNDSRTPVYAQHHGRRLVGNRRNRRHCNAVTPCRTVCRDDMNACRTGCH 70 80 90 100 110 120. 130 140 150 RITERSLKIFLQIRHFSPLNCPLYKNAAHKASPHVQQFX m142.pep RITERSLKSFLQIRHFSPLNRPLYKNAAHKASPHVQQFX g142 130 140 The following partial DNA sequence was identified in N. meningitidis <SEQ ID 563>: a142.seq ATGCGTGCCG ATTTCATGTT TGCCGACAAT ATGCCCGTGC AGGTGCGCCA ACGCGCCCTC TATTTCAAGT TGTCCCGTTT TGCCGCGATG CCAGATGTGG 51 TAGGCAAACC GCTCTTCGGG CGACAGGCCG GTCAGCCCGG CAAAATGTTC 101 GGCAACATCC TGATGTTCGT CCGCCAGCGT ATTGATGCAG AGGCTGCCGT 201 TTTCCGACAG GATCGGAATG ATTCGCGCAC TCCGGTTGAT GCACAGCATC ACGGTCGGCG GCTCGTCCGT AACCGGCGCA ACCGCCGTCA TTGTAATGCC 251 GTAACGCCCT GCCGCACCGT CTGTCGTGAT GACATGAACG CCTGCCGCAC 301 351 AGGATGCCAT CGCATCACGG AACGAAGTTT GAAAAGTTTT CTGCAAATCC GCCATTTTTC CCCTTTAAAC TGTCCCCTAT ATAAGAATGC TGCACACAAG 401 451 GCACCCCCA TGTGCAGCAG TTCTGATTCA AAAAGCCGTC GGTCGGACAT 501 TTCCGCGCGT TACGGCGTAT TACGAGTTCA ACGCATCCTC GATTTTGGCA AGTTCTGCCA ACAGGTCTTT AAGCAGCAGC ATTTTCTCGC GGCCCAGCAC 551 601 TTCCTCGATA GCGTCGTAAC GCTCGTCCAC TTCTTCGCCG ATTTCCTCAT ACAGCTTCTC GCCCTCGGCA GTCAGCTTCA GAAAAACACG TCGTTGGTCG TTGGAAGGTT TCAGGCGGAC AACCAAACCC GCTTTTTCAA GGCGGGTCAG 701 GATACCGGTC AGGCTGGGGC GCAAAATGCA CGCCTGATTC GCCAAATCTT

BNSDOCID: <WO__9957280A2_J>

WO 99/57280



801 851 901	TGATCGGTAA TATTCGCCTG ATTCAGAATA GACGGATAAT CCGCCATTGC GGCTTCCCTT GCCTGTATCA GACCGATATT GATAGACGCA TGTTTTGA	
This correspond	ds to the amino acid sequence <seq 142.a="" 564;="" id="" orf="">:</seq>	
a142.pep		
1 51 101 151 201 251 301	MRADFMFADN MPVQVRQRAL YFKLSRFAAM PDVVGKPLFG RQAGQPGKMF GNILMFVRQR IDAEAAVFRQ DRNDSRTPVD AQHHGRRLVR NRRNRRHCNA VTPCRTVCRD DMNACRTGCH RITERSLKSF LQIRHFSPLN CPLYKNAAHK APPMCSSSDS KSRRSDISAR YGVLRVQRIL DFGKFCQQVF KQQHFLAAQH FLDSVVTLVH FFADFLIQLL ALGSQLQKNT SLVVGRFQAD NQTRFFKAGQ DTGQAGAQNA RLIRQILKVQ RAVFRQKTDN PPLLIGNIRL IQNRPELGHQ GFPCLYQTDI DRRMF*	
m142/a142 96	.1% identity in 153 aa overlap	
m142.pep	10 20 30	_
	*** CALCADAME VOVEDBA! VEVI CD CA AMBRET	U
a142		ı
	10 20 20 20 20 20 20 20 20 20 20 20 20 20	Ŕ
	20 30 40 50 60	
m142.pep	70 80 90 100 110 120	_
ma42.pep	IDALAAVERODRNDSRTPVDAOUUCRDI VOVDRADA	)
a142		
	70 20 20 20 20 20 20 20 20 20 20 20 20 20	1
	70 80 90 100 110 120	
	130 140 150 159	
m142.pep	RITERSLKIFLOIRHFSPINCPLYKNAAUVAGDUVGGT	
a142		
uliz	RITERSLKSFLQIRHFSPLNCPLYKNAAHKAPPMCSSSDSKSRRSDISARYGVLRVQRIL	
	130 140 150 160 170 180	
a142		
	DFGKFCQQVFKQQHFLAAQHFLDSVVTLVHFFADFLIQLLALGSQLQKNTSLVVGRFQAD  190 200 210 220 220	
	200 210 220 230 240	

# The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 565>:

9-10.0Cq				_
1 51 101 151 201 251 301 351 401 451 501 651 701 751	ATGTTGAGCT TCGGCTATCT CTCGCAGATG AGCCGCATTT TGGGCTGGTT TTTCATCCTG ATAGTGGGCT ACTACTCAGA CCGCCTGCCG TATCTGCTTT TTTTGATGCC GAACTCGGGC GCCTTGTCGT TCGGCGCGCTT AGCAGAAAAG CTACGCCTAC GCGGTTGTGG CAGCAATTCT GAACACTGCC GAGAAAGGCC ATGTGGGTGC GGCGTTACTG GTCAAAGAAT ACGACCCGGA CGCCCGGAAT CGCCCGGAAT CTAAAGTGTT TTGGACGGTT TTCCGGTATA	TTCAAACGCT CCGCCCCCTGC CCGCCCCCTGCT CCGCACTTGCCT CGATGATTGCCT CGATGATTGCCTTTGTGCGATTTGTGCGATTTGTGCCACAACTACCACAACTGGTT ACTCCGCTACCAACTGGTTACCCCCCAACTGGTTACCCCCCCAACTGGTTACCCCCCCAACTGGTTACCCCCCCC	GCGGGGATGCT AAGCCGCGCT CGATTGCACG CTGTTGGACG CGGCGATATG GTTTCTTAGC TTCGCGTATA AACCGTGGTC GTGCGTTCAC CGTTACCACG CGAACTCTTA AGTTTTTCTG	CCGCACAATT GGTTCAGCCG TGGGCGGCCG ATCGTGATGA GTCGCTGGCG GTCAACGAGG GAATACGACC TCGGTTTGGC GTAGCATTCT AATCTCCAAA GCATCGATGT
		AGATGATGGT	CGGCGATATG	
		GCCGTTTCTC		O- I - I - I - I - I - I - I - I - I - I
	GAACACTGCC GAGAAAGGCG			
651				
701		4 44 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4		
751	TTCCGGTATA TGTGGACTTA	CTCCCCACCC		
801	GCACACTACC GATGCGTCTT			AAAACGTCTG
851	ACGGCGTTTT GGCGGCGGTG	CCGTAGGCCA		GGCAACCGGT
901	ATTCTGGCAA AAGTACCGAA	TAGTCGGTTG	CGGCGGTGAT	TTGTTCGTTT
951	TTTGGCTTTG GGCGCGCTCG	TAAATACCAT	AAGGCGGGTT	ATTTCGGCTG
1001	0000000100	GTTTCTTCTC	TATCTTCTTC	ATCTACAATC
1051		TATATCTTAA		TTGGGCGGGC
1101	CAMCCAMACE	GATTGTGGCC	AACGCTTTGT	CGGGCAAACA
1151		TGTttaacgg	CtctatCTGT	ATGCcgcaaa
	tcgTcgctTC GctgttgAGT	TTCGTGCTTT	TCCCGATGCT	GGGCGGCCAT

```
1201 CAGGCAACCA TGTTCTTGGT TGCAGGCGCA GTCTTGCTGC TGGGAGCCTT 1251 CTCAGTCTGT CTGATTAAAG AGATCCACGG CGGGGTTTGA
```

#### This corresponds to the amino acid sequence <SEQ ID 566; ORF 143.ng>:

```
g143.pep

1 MLSFGYLGVQ TAFTLQSSQM SRIFQTLGAD PHNLGWFFIL PPLAGMLVQP
51 IVGYYSDRTW KPRLGGRRLP YLLYGTLIAV IVMILMPNSG SFGFGYASLA
101 ALSFGALMIA LLDVSSNMAM QPFKMMVGDM VNEEQKSYAY GIQSFLANTD
151 AVVAAILPFV FAYIGLANTA EKGVVPQTVV VAFYVGAALL IITSAFTISK
201 VKEYDPETYA RYHGIDVAAN QEKANWFELL KTAPKVFWTV TPVQFFCWFA
251 FRYMWTYSAG AIAENVWHTT DASSVGHQEA GNRYGVLAAV *SVAAVICSF
301 ILAKVPNKYH KAGYFGCLAL GALGFFSIFF IYNQYALILS YILIGIAWAG
```

351 IITYPLTIVA NALSGKHMDT YLGLFNGSVC MPQIVASLLS FVLFPMLGGH 401 QATMFLVAGA VLLLGAFSVC LIKEIHGGV*

#### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 567>:

```
m143.seq
          ATGCTCAGTT TCGGCTTTCT CGGCGTTCAG ACGGCCTTTA CCCTGCAAAG
      51
          CTCGCAAATG AGCCGCATTT TTCAAACGCT AGGCGCAGAC CCGCACAATT
         TGGGCTGGTT TTTCATCCTG CCGCCGCTGG CGGGGATGCT GGTGCAGCCG
     101
     151 ATTGTCGGCC ATTACTCCGA CCGCACTTGG AAGCCGCGTT TGGGCGGCCG
     201 CCGTCTGCCG TATCTGCTTT ATGGCACGCT GATTGCGGTT ATTGTGATGA
     251 TTTTGATGCC GAACTCGGGC AGCTTCGGTT TCGGCTATGC GTCGCTGGCG
     301 GCTTTGTCGT TCGGCGCGCT GATGATTGCG CTGTTAGACG TGTCGTCAAA
         TATGGCGATG CAGCCGTTTA AGATGATGGT CGGCGACATG GTCAACGAGG
     351
     401 AGCAGAAAGG CTACGCCTAC GGGATTCAAA GTTTCTTAGC AAATACGGGC
     451 GCGGTCGTGG CGGCGATTCT GCCGTTTGTG TTTGCGTATA TCGGTTTGGC
    501 GAACACCGCC GAGAAAGGCG TTGTGCCGCA GACCGTGGTC GTGGCGTTTT
    551 ATGTGGGTGC GGCGTTGCTG GTGATTACCA GCGCGTTCAC GATTTTCAAA
    601 GTGAAGGAAT ACGATCCGGA AACCTACGCC CGTTACCACG GCATCGATGT
    651 CGCCGCGAAT CAGGAAAAAG CCAACTGGAT CGAACTCTTG AAAACCGCGC
    701 CTAAGGCGTT TTGGACGGTT ACTTTGGTGC AATTCTTCTG CTGGTTCGCC
    751 TTCCAATATA TGTGGACTTA CTCGGCAGGC GCGATTGCGG AAAACGTCTG
    801 GCACACCACC GATGCGTCTT CCGTAGGTTA TCAGGAGGCG GGTAACTGGT
    851 ACGGCGTTTT GGCGGCGGTG CAGTCGGTTG CGGCGGTGAT TTGTTCGTTT
         GTATTGGCGA AAGTGCCGAA TAAATACCAT AAGGCGGGTT ATTTCGGCTG
    901
    951 TTTGGCTTTG GGCGCGCTCG GCTTTTTCTC CGTTTTCTTC ATCGGCAACC
   1001 AATACGCGCT GGTGTTGTCT TATACCTTAA TCGGCATCGC TTGGGCGGGC
   1051 ATTATCACTT ATCCGCTGAC GATTGTGACC AACGCCTTGT CGGGCAAGCA
   1101 TATGGGCACT TACTTGGGCT TGTTTAACGG CTCTATCTGT ATGCCTCAAA
   1151
         TCGTCGCTTC GCTGTTGAGT TTCGTGCTTT TCCCTATGCT GGGCGGCTTG
   1201 CAGGCCACTA TGTTCTTGGT AGGGGGCGTC GTCCTGCTGC TGGGCGCGTT
   1251 TTCCGTGTTC CTGATTAAAG AAACACACGG CGGGGTTTGA
```

#### This corresponds to the amino acid sequence <SEQ ID 568; ORF 143>:

```
m143.pep

1 MLSFGFLGVQ TAFTLQSSQM SRIFQTLGAD PHNLGWFFIL PPLAGMLVQP
51 IVGHYSDRTW KPRLGGRRLP YLLYGTLIAV IVMILMPNSG SFGFGYASLA
101 ALSFGALMIA LLDVSSNMAM QPFKMWGDM VNEEQKGYAY GIQSFLANTG
151 AVVAAILPFV FAYIGLANTA EKGVVPQTVV VAFYVGAALL VITSAFTIFK
201 VKEYDPETYA RYHGIDVAAN QEKANWIELL KTAPKAFWTV TLVQFFCWFA
251 FQYMWTYSAG AIAENVWHTT DASSVGYQEA GNWYGVLAAV QSVAAVICSF
301 VLAKVPNKYH KAGYFGCLAL GALGFFSVFF IGNQYALVLS YTLIGIAWAG
351 IITYPLTIVT NALSGKHMGT YLGLFNGSIC MPQIVASLLS FVLFPMLGGL
401 QATMFLVGGV VLLLGAFSVF LIKETHGGV*
```

# Computer analysis of this amino acid sequence gave the following results: Homology with a predicted ORF from N. gonorrhoeae

```
m143 / g143 93.9% identity in 429 aa overlap
```

10 20 30 40 50 60

m143.pep	MLSFGFLGVQTAF	TLOSSOMSRI	FOTLGADPHI	NLGWFFILPPL	AGMLVQPIV	GHYSDRTW
g143		3 1 1 1 1 1 1 1 1 1				
9110	MLSFGYLGVQTAF	TLQSSQMSRI 20	POTLGADPHI	NLGWFFILPPL	AGMLVQPIV	GYYSDRTW
	10	20	30	40	50	60
	70	80	90	100	110	
m143.pep	KPRLGGRRLPYLL	YGTLIAVIVM	ILIMPNSGSEG	TECVACIANT O	110 FCALMIALL	120
1.10		1 1 1 1 1 1 1 1 1 1	11111111	111111111		
g143		TGITIMATAM	ILMPNSGSFG	FGYASLAALS	FGALMIALLI	OVSSNMAM
	70	80	90	100	110	120
	130	140				
m143.pep		140 FORGVAVCTO	150	160	170	180
	QPFKMMVGDMVNEI                  QPFKMMVGDMVNEI		SELANTGAVV	'AAILPFVFAY	IGLANTAEKG	VVPQTVV
g143	QPFKMMVGDMVNEI	EOKSYAYGIO	SFLANTDAVA			
	130	140	150	160	IGLANTAEKG 170	
				100	170	180
-142	190	200	210	220	230	240
m143.pep	VAFYVGAALLVITS	AFTIFKVKE	YDPETYARYH	GIDVAANQEKA	MUTELTEE	Date
g143						
9410	VAFYVGAALLIITS	AFTISKVKE 200	LDEELIARAH	GIDVAANQEKA	NWFELLKTA	PKVFWTV
	150	200	210	220	230	240
	250	260	270	280	200	
m143.pep	TLVQFFCWFAFQYM	WTYSAGATAR	מת ייים אות אות אות אות	CUCVOE B CARRY	290	300
4.40				1 [ ] • ] [ ] 1 ] 1	111111 11	
g143		H T T DYGYTAE	CNVWHTTDASS	SVGHQEAGNRY	GVLAAVXSV	AAVICSE
	250	260	270	280	290	300
	310	320				
m143.pep			330	340	350	360
	VLAKVPNKYHKAGY		FESVEETGNC	QYALVLSYTLI:	GIAWAGIIT	YPLTIVT
g143	ILAKVPNKYHKAGY	FGCLALGALG	FFSIFFTYNC			:
	310	320	330	340	350	360
					330	360
m143.pep	370	380	390	400	410	420
mr43.pep	NALSGKHMGTYLGL	NGSICMPQI	VASLLSFVLF	PMLGGLQATM	FLVGGVVLLI	GAFSVF
g143			4		111-1-11:	
-	NALSGKHMDTYLGLI	380	VASLLSFVLF 390	PMLGGHQATMI	FLVAGAVLLL	GAFSVC
		300	390	400	410	420
	430					
m143.pep	LIKETHGGVX	•				
~1.42						
g143	LIKEIHGGVX					
	430					

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 569>:

al43.seq				J	
1	ATGCTCAGTT	TCGGCTTTCT	CGGCGTTCAG	ACGGCCTTTA	CCCTGCAAAG
51	CTCGCAGATG	AGCCGCATCT	TCCAGACGCT	CGGTGCCGAT	
101	TCGGCTGGTT	CTTTATCCTG		CGGGGATGCT	
151	ATTGTCGGCC	ATTACTCCGA	CCGCACTTGG	AACCCCCCCC	
201	CCGTCTGCCG	TATCTGCTTT	ATGGCACGCT	CAMMCCCCCM	TGGGCGGCCG
251	TTTTGATGCC	GAACTCGGGC	AGCTTCGGTT	TCCCCTATION	ATTGTGATGA
301	GCTTTGTCGT	TCGGCGCGCT	CATCATTCCC	CECETATGC	GTCGCTGGCG
351	TATGGCGATG	CAGCCGTTTA			
401	AGCAGAAAGG	CTACGCCTAC	CCCAMMCAAA	CGGCGACATG	GTCAACGAGG
451		CTACGCCTAC	CCCCTTTCAAA	GTTTCTTAGC	GAATACGGGC
501		CGGCGATTCT	TTCTCTC	TTTGCGTATA	TCGGTTTGGC
551		GAGAAAGGCG	CHEARTH	GACCGTGGTC	GTGGCGTTTT
601		GGCGTTGCTG	AAGGERAGE	GCGCGTTCAC	GATTTTCAAA
651		ACAATCCGGA	CCRRCTACGCC	CGTTACCACG	GCATCGATGT
701		CAGGAAAAAG	CCAACTGGAT	CGAACTCTTG	AAAACCGCGC
		TTGGACGGTT	ACTITGGTGC	AATTCTTCTC	CTGGTTCGCC

801 851 901 951	GCACACCACC ACGGCGTTTT GTATTGGCGA TTTGGCTTTG	GATGCGTCTT GGCGGCGGTG AAGTGCCGAA GGCGCGCTCG	CTCGGCAGGC CCGTAGGTTA CAGTCGGTTG TAAATACCAT GCTTTTTCTC	TCAGGAGGCG CGGCGGTGAT AAGGCGGGTT CGTTTTCTTC	GGTAACTGGT TTGTTCGTTT ATTTCGGCTG ATCGGCAACC	
1001 1051			TATACCTTAA GATTGTGACC			
1101			TGTTTAACGG			
1151			TTCGTGCTTT			
1201	CAGGCCACTA	TGTTCTTGGT	AGGGGGCGTC	GTCCTGCTGC		
1251	TTCCGTGTTC	CTGATTAAAG	AAACACACGG	CGGGGTTTGA		
This correspond	s to the amin	o acid seque	ence <seq i<="" td=""><td>D 570; ORF</td><td>143.a&gt;:</td><td></td></seq>	D 570; ORF	143.a>:	
a143.pep 1	MI.SEGELGVO	TAFTLOSSOM	SRIFQTLGAD	PHSLGWFFTI.	PPT.ACMT.VOP	
51			YLLYGTLIAV			
101	ALSFGALMIA	LLDVSSNMAM	QPFKMMVGDM	VNEEQKGYAY	GIQSFLANTG	
151	AVVAAILPFV	FAYIGLANTA	EKGVVPQT <u>VV</u>	VAFYVGAALL	VITSAFTIFK	
201 251			QEKANWIELL			
301			DASSVGYQEA GALGFFSVFF			
351			YLGLFNGSIC			
401		VLLLGAFSVF				
m143/a143	99.5% i	dentity in	429 aa over	lap		
•		10	20 3	0 40	50	60
m143.pep					PPLAGMLVQPIV	
4.40					1111111111	
a143	MLSFGFL		QMSRIFQTLGA 20 3:		PPLAGMLVQPIV 50	GHYSDRTW 60
		10	20 3	0 40	20	80
		70	80 9	0 100	110	120
m143.pep	KPRLGGRI	RLPYLLYGTLI	AVIVMILMPNS	GSFGFGYASLA	ALSFGALMIALI	DVSSNMAM
-142					IIIIIIIIIIII	
<b>a1</b> 43		RLPYLLYGTLI	AVIVMILMPNS	GSFGFGYASLA	ALSFGALMIALI	DVSSNMAM
<b>a1</b> 43		RLPYLLYGTLI		GSFGFGYASLA		
	KPRLGGRI	RLPYLLYGTLI 70 130 1	AVIVMILMPNS BO 9 40 15	GSFGFGYASLAI 0 100 0 160	ALSFGALMIALI 110 170	DVSSNMAM 120. 180 [°]
a143 m143.pep	KPRLGGRI : QPFKMMV0	RLPYLLYGTLI 70 130 1 GDMVNEEQKGY	AVIVMILMPNS 80 9 40 15 AYGIQSFLANT	GSFGFGYASLA 0 100 0 160 GAVVAAILPFV	ALSFGALMIALI 110 170 FAYIGLANTAEK	DVSSNMAM 120. 180° GVVPQTVV
m143.pep	KPRLGGRI QPFKMMV0	RLPYLLYGTLI 70 130 1 GDMVNEEQKGY	AVIVMILMPNS 80 9: 40 15: AYGIQSFLANT	GSFGFGYASLAI 0 100 0 160 GAVVAAILPFVI	ALSFGALMIALI 110 170 FAYIGLANTAEK	DVSSNMAM 120. 180° GVVPQTVV
	KPRLGGRI QPFKMMVO         QPFKMMVO	RLPYLLYGTLI 70 130 1 GDMVNEEQKGY           GDMVNEEQKGY	AVIVMILMPNS 80 9: 40 15: AYGIQSFLANT	GSFGFGYASLAI 0 100 0 160 GAVVAAILPFVI           GAVVAAILPFVI	ALSFGALMIALI 110 170 FAYIGLANTAEK	DVSSNMAM 120. 180° GVVPQTVV
m143.pep	KPRLGGRI QPFKMMVO         QPFKMMVO	RLPYLLYGTLI 70 130 1 GDMVNEEQKGY            GDMVNEEQKGY 130 1	AVIVMILMPNS 80 9 40 15 AYGIQSFLANT          AYGIQSFLANT 40 15	GSFGFGYASLAI 0 100 0 160 GAVVAAILPFVI            GAVVAAILPFVI 0 160	ALSFGALMIALI 110 170 FAYIGLANTAEK           FAYIGLANTAEK	DVSSNMAM 120. 180° GVVPQTVV         GVVPQTVV
m143.pep a143	KPRLGGRI QPFKMMV        QPFKMMV	RLPYLLYGTLI 70  130  130  GDMVNEEQKGY             GDMVNEEQKGY  130  1	AVIVMILMPNS6 80 96 40 15 AYGIQSFLANT6             AYGIQSFLANT6 40 15	GSFGFGYASLAM 0 100 0 160 GAVVAAILPFV	ALSFGALMIALI 110 170 FAYIGLANTAEK           FAYIGLANTAEK 170 230	DVSSNMAM 120. 180  GGVVPQTVV                   GGVVPQTVV 180 240
m143.pep	KPRLGGRI QPFKMMV        QPFKMMV VAFYVGA	RLPYLLYGTLI 70  130  130  GDMVNEEQKGY             GDMVNEEQKGY  130  1  190  2  ALLVITSAFTI	AVIVMILMPNS6 80 96 40 156 AYGIQSFLANT6            AYGIQSFLANT6 40 156 00 21 FKVKEYDPETY	GSFGFGYASLAM 0 100 0 160 GAVVAAILPFVI            GAVVAAILPFVI 0 160 0 220 ARYHGIDVAAN	ALSFGALMIALI 110 170 FAYIGLANTAEK            FAYIGLANTAEK 170 230 DEKANWIELLKT	DVSSNMAM 120. 180°:  GVVPQTVV                    GVVPQTVV 180 240 CAPKAFWTV
m143.pep a143 m143.pep	KPRLGGRI QPFKMMV         QPFKMMV VAFYVGA	RLPYLLYGTLI. 70  130  130  1 GDMVNEEQKGY.	AVIVMILMPNS6 80 96 40 156 AYGIQSFLANT6            AYGIQSFLANT6 40 156 00 216 FKVKEYDPETY	GSFGFGYASLAM 0 100 0 160 GAVVAAILPFVI                         GAVVAAILPFVI 0 160 0 220 ARYHGIDVAAN	ALSFGALMIALI 110 170 FAYIGLANTAEK            FAYIGLANTAEK 170 230 DEKANWIELLKT	DVSSNMAM 120. 180°: GVVPQTVV                   GVVPQTVV 180 240 CAPKAFWTV
m143.pep a143	VAFYVGAL	RLPYLLYGTLI 70  130  1 GDMVNEEQKGY                         GDMVNEEQKGY  130  1  190  2  ALLVITSAFTI                       ALLVITSAFTI	AVIVMILMPNS6 80 96 40 156 AYGIQSFLANT6            AYGIQSFLANT6 40 156 00 216 FKVKEYDPETY	GSFGFGYASLAM 0 100 0 160 GAVVAAILPFVI            GAVVAAILPFVI 0 160 0 220 ARYHGIDVAAN	ALSFGALMIALI 110 170 FAYIGLANTAEK            FAYIGLANTAEK 170 230 DEKANWIELLKT	DVSSNMAM 120. 180°: GVVPQTVV                   GVVPQTVV 180 240 CAPKAFWTV
m143.pep a143 m143.pep	QPFKMMV()	RLPYLLYGTLI 70  130 1 GDMVNEEQKGY            GDMVNEEQKGY 130 1  190 2 ALLVITSAFTI          ALLVITSAFTI 190 2	AVIVMILMPNS6 80 9 40 15 AYGIQSFLANT0                     AYGIQSFLANT0 40 15 00 21 FKVKEYDPETY.                   FKVKEYNPETY. 00 21	GSFGFGYASLAM 0 100 0 160 GAVVAAILPFV 0 160 0 220 ARYHGIDVAAN 0 220 ARYHGIDVAAN 0 220	ALSFGALMIALI 110 170 FAYIGLANTAEK            FAYIGLANTAEK 170 230 DEKANWIELLKT	DVSSNMAM 120. 180 CGVVPQTVV         CGVVPQTVV 180 240 CAPKAFWTV
m143.pep a143 m143.pep a143	VAFYVGA	RLPYLLYGTLI 70  130 1 GDMVNEEQKGY            GDMVNEEQKGY 130 1  190 2 ALLVITSAFTI          ALLVITSAFTI 191 2 250 2	AVIVMILMPNS6 80 9 40 15 AYGIQSFLANT6	GSFGFGYASLAM 0 100 0 160 GAVVAAILPFVI 0 160 0 220 ARYHGIDVAAN 1	ALSFGALMIALI 110 170 FAYIGLANTAEK             FAYIGLANTAEK 170 230 QEKANWIELLKT            QEKANWIELLKT 230 290	.DVSSNMAM 120. 180 CGVVPQTVV          CGVVPQTVV 180 240 CAPKAFWTV         CAPKAFWTV 240 300
m143.pep a143 m143.pep	VAFYVGAL	RLPYLLYGTLI 70  130 1 GDMVNEEQKGY            GDMVNEEQKGY 130 1  190 2 ALLVITSAFTI          ALLVITSAFTI 191 2  250 2 WFAFQYMWTYS	AVIVMILMPNS 80 9 40 15 AYGIQSFLANT                     AYGIQSFLANT 40 15 00 21 FKVKEYDPETY                     FKVKEYNPETY 00 21 60 27 AGGAIAENVWHT	GSFGFGYASLAM 0 100 0 160 GAVVAAILPFVI 0 160 0 220 ARYHGIDVAAN 1	ALSFGALMIALI 110  170  FAYIGLANTAEK            FAYIGLANTAEK 170  230  QEKANWIELLKT            QEKANWIELLKT 230  290  GNWYGVLAAVQS	DVSSNMAM 120. 180 GVVPQTVV IIIIIIII GVVPQTVV 180 240 CAPKAFWTV IIIIIIIII CAPKAFWTV 240 300 GVAAVICSF
m143.pep a143 m143.pep a143	VAFYVGAL VAFYVGAL VAFYVGAL VAFYVGAL	RLPYLLYGTLI 70  130 1 GDMVNEEQKGY            GDMVNEEQKGY 130 1  190 2 ALLVITSAFTI           ALLVITSAFTI 191 2  250 2 WFAFQYMWTYS	AVIVMILMPNS6 80 99 40 15 AYGIQSFLANT0 1                       AYGIQSFLANT0 40 15 00 21 FKVKEYDPETY FKVKEYDPETY FKVKEYNPETY 00 21 60 27 AGAIAENVWHT	GSFGFGYASLAM 0 100 0 160 GAVVAAILPFVI 0 160 0 220 ARYHGIDVAAN 0 220 4RYHGIDVAAN 0 220 0 280 TDASSVGYQEA	ALSFGALMIALI 110  170  FAYIGLANTAEK            FAYIGLANTAEK 170  230  QEKANWIELLKT            QEKANWIELLKT 230  290  GNWYGVLAAVQS	180° CGVVPQTVV                 CGVVPQTVV   180  240 CAPKAFWTV                 CAPKAFWTV   240  300 CVAAVICSF
m143.pep a143 m143.pep a143	VAFYVGAL VAFYVGAL VAFYVGAL VAFYVGAL VAFYVGAL VAFYVGAL VAFYVGAL	RLPYLLYGTLI 70  130 1 GDMVNEEQKGY            GDMVNEEQKGY 130 1  190 2 ALLVITSAFTI           ALLVITSAFTI  190 2  SALLVITSAFTI  190 2  WFAFQYMWTYS           WFAFQYMWTYS	AVIVMILMPNS6 80 99 40 15 AYGIQSFLANT0 1                       AYGIQSFLANT0 40 15 00 21 FKVKEYDPETY FKVKEYDPETY FKVKEYNPETY 00 21 60 27 AGAIAENVWHT	GSFGFGYASLAM 0 100 0 160 GAVVAAILPFV	ALSFGALMIALI 110  170  FAYIGLANTAEK            FAYIGLANTAEK 170  230  QEKANWIELLKT            QEKANWIELLKT 230  290  GNWYGVLAAVQS	180° CGVVPQTVV                 CGVVPQTVV   180  240 CAPKAFWTV                 CAPKAFWTV   240  300 CVAAVICSF
m143.pep a143 m143.pep a143	QPFKMMVO                 QPFKMMVO                 QPFKMMVO                 VAFYVGA                 VAFYVGA                 TLVQFFC	RLPYLLYGTLI 70  130 1 GDMVNEEQKGY             GDMVNEEQKGY 130 1  190 2 ALLVITSAFTI           ALLVITSAFTI 190 2  250 2 WFAFQYMWTYS           WFAFQYMWTYS 250 2	AVIVMILMPNS6 80 9 40 15 AYGIQSFLANT6            AYGIQSFLANT4 40 15 00 21 FKVKEYDPETY           FKVKEYNPETY 00 21 60 27 AGAIAENVWHT          AGAIAENVWHT	GSFGFGYASLAM 0 100 0 160 GAVVAAILPFVI                         GAVVAAILPFVI 0 160 0 220 ARYHGIDVAAN 0 220 0 280 TDASSVGYQEA                         TDASSVGYQEA 0 280	ALSFGALMIALI 110  170  FAYIGLANTAEK             FAYIGLANTAEK 170  230  QEKANWIELLKT            QEKANWIELLKT 230  290  GNWYGVLAAVQS              GNWYGVLAAVQS 290	DVSSNMAM 120.  180 GVVPQTVV 1       GVVPQTVV 180 240 CAPKAFWTV 240 SVAAVICSF         GVAAVICSF
m143.pep a143 m143.pep a143 m143.pep a143	QPFKMMVO                 QPFKMMVO                 QPFKMMVO                 VAFYVGAI               VAFYVGAI               TLVQFFC	RLPYLLYGTLI 70  130 1 GDMVNEEQKGY             GDMVNEEQKGY 130 1  190 2 ALLVITSAFTI           ALLVITSAFTI 190 2  250 2 WFAFQYMWTYS           WFAFQYMWTYS 250 2	AVIVMILMPNS6 80 99 40 15 AYGIQSFLANT6             AYGIQSFLANT4 40 15 00 21 FKVKEYDPETY           FKVKEYNPETY 00 21 60 27 AGAIAENVWHT          AGAIAENVWHT 60 27	GSFGFGYASLAM 0 100 0 160 GAVVAAILPFV	ALSFGALMIALI 110  170  FAYIGLANTAEK             FAYIGLANTAEK 170  230  QEKANWIELLKT 230  290  GNWYGVLAAVQS              GNWYGVLAAVQS 290 350	DVSSNMAM 120. 180 GVVPQTVV 11111111 GVVPQTVV 180 240 CAPKAFWTV 1111111 CAPKAFWTV 240 300 GVAAVICSF 11111111 GVAAVICSF 300 360
m143.pep a143 m143.pep a143	QPFKMMVO                   QPFKMMVO   VAFYVGAI                 VAFYVGAI   TLVQFFCI                 TLVQFFCI	RLPYLLYGTLI. 70  130 1 GDMVNEEQKGY.             GDMVNEEQKGY. 130 1  190 2 ALLVITSAFTI           ALLVITSAFTI 190 2  EXECUTE OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF THE STREET OF TH	AVIVMILMPNS6 80 99 40 15 AYGIQSFLANT6                         AYGIQSFLANT6 40 15 00 21 FKVKEYDPETY.                         FKVKEYNPETY. 00 21 60 27 AGAIAENVWHT                       AGAIAENVWHT 60 27 20 33 ALGALGFFSVF	GSFGFGYASLAM 0 100 0 160 GAVVAAILPFV	ALSFGALMIALI 110  170  FAYIGLANTAER              FAYIGLANTAER              EAYIGLANTAER 170  230  QEKANWIELLKT 230  290  GNWYGVLAAVQS	DVSSNMAM 120. 180° CGVVPQTVV 11111111 CGVVPQTVV 180 240 CAPKAFWTV 11111111 CAPKAFWTV 240 300 CVAAVICSF 11111111 CVAAVICSF 300 360 CTYPLTIVT
m143.pep a143 m143.pep a143 m143.pep a143	QPFKMMV()	RLPYLLYGTLI 70  130 1  GDMVNEEQKGY              GDMVNEEQKGY  130 1  190 2  ALLVITSAFTI             ALLVITSAFTI              ALLVITSAFTI               WFAFQYMWTYS                WFAFQYMWTYS  250 2  310 3  KYHKAGYFGCL	AVIVMILMPNS 80 9 40 15 AYGIQSFLANT                     AYGIQSFLANT 40 15 00 21 FKVKEYDPETY                   FKVKEYNPETY 00 21 AGAIAENVWHT                 AGAIAENVWHT 60 27 20 33 ALGALGFFSVF	GSFGFGYASLAM 0 100 0 160 GAVVAAILPFV 0 160 0 220 ARYHGIDVAAN 0 220 0 280 TDASSVGYQEA 0 280 0 280 0 280 0 340 FIGNQYALVLS	ALSFGALMIALI 110  170  FAYIGLANTAEK            FAYIGLANTAEK 170  230  DEKANWIELLKT 230  ZEKANWIELLKT 230  GNWYGVLAAVQS 1           GNWYGVLAAVQS 290  350  YTLIGIAWAGII	DVSSNMAM 120. 180 CGVVPQTVV 11111111 CGVVPQTVV 180 240 CAPKAFWTV 241 CAPKAFWTV 240 SVAAVICSF 300 SVAAVICSF 300 TYPLTIVT
m143.pep a143 m143.pep a143 m143.pep a143	VAFYVGAL  VAFYVGAL  VAFYVGAL  VAFYVGAL  VAFYVGAL  VAFYVGAL  VLAKVPNI  VLAKVPNI  VLAKVPNI	RLPYLLYGTLI 70  130 1 GDMVNEEQKGY             GDMVNEEQKGY 130 1  190 2 ALLVITSAFTI            ALLVITSAFTI  1         WFAFQYMWTYS             WFAFQYMWTYS 250 2 WFAFQYMWTYS             WFAFQYMWTYS 1          WFAFQYMWTYS 250 2	AVIVMILMPNS 80 9 40 15 AYGIQSFLANT                     AYGIQSFLANT 40 15 00 21 FKVKEYDPETY                   FKVKEYNPETY 00 21 AGAIAENVWHT                 AGAIAENVWHT 60 27 20 33 ALGALGFFSVF	GSFGFGYASLAM 0 100 0 160 GAVVAAILPFV	ALSFGALMIALI 110  170  FAYIGLANTAER              FAYIGLANTAER              EAYIGLANTAER 170  230  QEKANWIELLKT 230  290  GNWYGVLAAVQS	DVSSNMAM 120. 180 CGVVPQTVV 11111111 CGVVPQTVV 180 240 CAPKAFWTV 241 CAPKAFWTV 240 SVAAVICSF 300 SVAAVICSF 300 TYPLTIVT
m143.pep a143 m143.pep a143 m143.pep a143	QPFKMMV()         QPFKMMV() VAFYVGAI         VAFYVGAI TLVQFFC()         TLVQFFC() VLAKVPN() VLAKVPN()	RLPYLLYGTLI 70  130 1 GDMVNEEQKGY.            GDMVNEEQKGY. 130 1  190 2 ALLVITSAFTI           ALLVITSAFTI            ALLVITSAFTI            WFAFQYMWTYS. 250 2 WFAFQYMWTYS. 250 2  310 3 KYHKAGYFGCL            KYHKAGYFGCL 310 3	AVIVMILMPNS 80 9 40 15 AYGIQSFLANT                     AYGIQSFLANT 40 15 00 21 FKVKEYDPETY                   FKVKEYNPETY 00 21 AGAIAENVWHT                 AGAIAENVWHT 60 27 20 33 ALGALGFFSVF                   ALGALGFFSVF 20 33	GSFGFGYASLAM 0 100 0 160 GAVVAAILPFV 0 160 0 220 ARYHGIDVAAN 0 220 0 280 TDASSVGYQEA 0 280 TDASSVGYQEA 0 280 0 340 FIGNQYALVLS 0 340	ALSFGALMIALI 110  170  FAYIGLANTAEK            FAYIGLANTAEK 170  230  DEKANWIELLKT            DEKANWIELLKT 230  350  YTLIGIAWAGII            YTLIGIAWAGII 350	180° CGVVPQTVV         CGVVPQTVV          CGVVPQTVV          CAPKAFWTV          CAPKAFWTV          CAPKAFWTV          CAPKAFWTV          CAPKAFWTV          CAPKAFWTV          CAPKAFWTV          CAPKAFWTV           CAPKAFWTV           CAPKAFWTV           CAPKAFWTV           CAPKAFWTV           CAPKAFWTV           CAPKAFWTV           CAPKAFWTV           CAPKAFWTV           CAPKAFWTV           CAPKAFWTV           CAPKAFWTV           CAPKAFWTV           CAPKAFWTV           CAPKAFWTV           CAPKAFWTV           CAPKAFWTV           CAPKAFWTV           CAPKAFWTV           CAPKAFWTV            CAPKAFWTV           CAPKAFWTV           CAPKAFWTV           CAPKAFWTV           CAPKAFWTV           CAPKAFWTV           CAPKAFWTV            CAPKAFWTV            CAPKAFWTV            CAPKAFWTV            CAPKAFWTV            CAPKAFWTV             CAPKAFWTV              CAPKAFWTV             CAPKAFWTV
m143.pep a143 m143.pep a143 m143.pep a143	QPFKMMV()         QPFKMMV() VAFYVGAL         VAFYVGAL TLVQFFCI         TLVQFFCI	RLPYLLYGTLI 70  130 1 GDMVNEEQKGY 11         GDMVNEEQKGY 130 1  190 2 ALLVITSAFTI ALLVITSAFTI ALLVITSAFTI 190 2 WFAFQYMWTYS 1         WFAFQYMWTYS 250 2  310 3 KYHKAGYFGCL            KYHKAGYFGCL            KYHKAGYFGCL            KYHKAGYFGCL 310 3	AVIVMILMPNS 80 9 40 15 AYGIQSFLANT                       AYGIQSFLANT 40 15 00 21 FKVKEYDPETY                     FKVKEYNPETY 00 21 60 27 AGAIAENVWHT   AGAIAENVWHT 60 27 20 33 ALGALGFFSVF                       ALGALGFFSVF                       ALGALGFFSVF                       ALGALGFFSVF                       ALGALGFFSVF                       ALGALGFFSVF                       ALGALGFFSVF                       ALGALGFFSVF                       ALGALGFFSVF                       ALGALGFFSVF                       ALGALGFFSVF                       ALGALGFFSVF                         ALGALGFFSVF                         ALGALGFFSVF                           ALGALGFFSVF                           ALGALGFFSVF                             ALGALGFFSVF                               ALGALGFFSVF                                     ALGALGFFSVF	GSFGFGYASLAM 0 100 0 160 GAVVAAILPFV 0 160 0 220 ARYHGIDVAAN 0 220 0 280 TDASSVGYQEA 0 280 TDASSVGYQEA 0 280 0 340 FIGNQYALVLS 0 340 0 340	ALSFGALMIALI 110  170  FAYIGLANTAER              FAYIGLANTAER               EAYIGLANTAER 170  230  QEKANWIELLKT 230  290  GNWYGVLAAVQS                GNWYGVLAAVQS 290  350  YTLIGIAWAGII	180° CGVVPQTVV                 CGVVPQTVV                 CGVVPQTVV   180  240 CAPKAFWTV               CAPKAFWTV 240  300 CVAAVICSF               CVAAVICSF               CVAAVICSF               CVAAVICSF               CVAAVICSF               CVAAVICSF               CVAAVICSF               CVAAVICSF               CVAAVICSF               CVAAVICSF                 CVAAVICSF                 CVAAVICSF                 CVAAVICSF                 CVAAVICSF                 CVAAVICSF                 CVAAVICSF                   CVAAVICSF                   CVAAVICSF                     CVAAVICSF                     CVAAVICSF                     CVAAVICSF                     CVAAVICSF                             CVAAVICSF

BNSDOCID: <WO__9957280A2_J_>

### The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 571>:

```
ATGAGCGATA CCCCCGCTAC CCGCGATTTC GGCCTGATCG ACGGGCGGGC
  1
 51
     CGTAACCGGC TATGTGCTGT CCAACCGGCG TGGTACGTGC GTCTTCGTGC
101
     TGGACTTGGG CGGGATTGTG CAGGAATTTT CCGTTTTGGC AGACGGCGTG
     CGCGAAAACC CCGTGGTGTC GTTCGACGAT GCGGCTTCCT ATGCGGACAA
151
    TCCGTTTCAG ATTAACAAGC AGATAGGGCG CGTGGCCGGA CGCATCCGCG
251 GTGCGGCGTT CGACATCAAC GGTAGGACTT ACCGCGTGGA GGCCAACGAA
301
    GGCAGGAACG CGCTGCACGG CGGTTCGCAC GGGCTGGCCG TTACCcgtTT
     CAACGCGGTG GCGGCAGACG GccgacggTt atCCCAACGA TTTGGatatT
351
    TCctaccgCT TGGACGAGGA CGGCCGGCTT ACCGTtaccT ATCGCGCCAC
401
    CGCgctCGGC GACACGGTGT TCGACCCGAC GCTGCACATT TACTGGCGGC
451
    TGGACGCGGG CCTGCACGAT GCGGTTCTGC ATATTCCGCA GGGCGGACAT
501
551
    ATTCCGGCCG ATGCCGAAAA ACTGCCCGTC TTAACGGTTT CAGACGGCCT
601
    CGAAGTATTT GA
```

#### This corresponds to the amino acid sequence <SEQ ID 572; ORF 144.ng>:

```
g144.pep

1 MSDTPATRDF GLIDGRAVTG YVLSNRRGTC VFVLDLGGIV QEFSVLADGV
51 RENPVVSFDD AASYADNPFQ INKQIGRVAG RIRGAAFDIN GRTYRVEANE
101 GRNALHGGSH GLAVTRFNAV AADGRRLSQR FGYFLPLGRG RPAYRYLSRH
151 RARRHGVRPD AAHLLAAGRG PARCGSAYSA GRTYSGRCRK TARLNGFRRP
201 RSI*
```

#### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 573>:

```
m144.seq
          ATGAGCGATA CCCCCGCTAC CCGCGATTTC GGTCTGATCG ACGGGCGTGC
       7
          CGTAACCGGC TATGTGCTGT CCAACCGGCG TGGTACGCGT GTCTGCGTGC
         TGGACTTGGG CGGGATTGTG CAGGAATTTT CCGTTTTGGC AGACGGCGTG
     101
         CGCGAAAACC TCGTGGTGTC GTTCGATGAT GCGGCTTCCT ATGCGGACAA
          TCCGTTTCAG ATTAACAAAC AGATAGGGCG CGTGGCCGGA CGCATCCGCG
     201
         GTGCGGCGTT CGACATCAAC GGCAGGACTT ACCGCGTGGA GGCCAACGAA
    251
         GGCAGGAACG CGCTGCACGG CGGTTCGCAC GGGCTGGCCG TTACCCGTTT
     351 CAACGCGGTG GCGGCAGACG GCCGTTCGGT GGTGCTGCGC AGCCGCCTGG
     401
         CAACAGTCGG CCGACGGTTA TCCCAACGAT TTGGATTTGG ATATTTCCTA
     451
         CCGCTTGGAC GAGGACGACC GGCTTACCGT TACCTATCGC GCCACCGCGC
         TCGGCGACAC GGTGTTCGAC CCGACGCTGC ACATTTACTG GCGGCTGGAC
    501
         GCGGGCCTGC ACGATGCGGT TCTGCATATT CCGCAGGGCG GACATATGCC
         GGCCGATGCC GAAAAACTGC CCGTCTCAAC GGTTTCAGAC GACCTCGAAG
    601
         TATTTGA
```

#### This corresponds to the amino acid sequence <SEQ ID 574; ORF 144>:

```
m144.pep

1 MSDTPATRDF GLIDGRAVTG YVLSNRRGTR VCVLDLGGIV QEFSVLADGV
51 RENLVVSFDD AASYADNPFQ INKQIGRVAG RIRGAAFDIN GRTYRVEANE
101 GRNALHGGSH GLAVTRFNAV AADGRSVVLR SRLATVGRRL SQRFGFGYFL
151 PLGRGRPAYR YLSRHRARRH GVRPDAAHLL AAGRGPARCG SAYSAGRTYA
201 GRCRKTARLN GFRRPRSI*
```

Computer analysis of this amino acid sequence gave the following results:

#### Homology with a predicted ORF from N. gonorrhoeae

```
m144 / g144 91.3% identity in 218 aa overlap
                                 20
                        10
                                                    40
                                          30
                                                             50
    m144.pep
                 MSDTPATRDFGLIDGRAVTGYVLSNRRGTRVCVLDLGGIVOEFSVLADGVRENLVVSFDD
                 {\tt MSDTPATRDFGLIDGRAVTGYVLSNRRGTCVFVLDLGGIVQEFSVLADGVRENPVVSFDD}
    g144
                                 20
                                          30
                                                    40
                                                            50
                                                                      60
                        70
                                 80
                                          90
                                                   100
                                                            110
                                                                     120
                 AASYADNPFQINKQIGRVAGRIRGAAFDINGRTYRVEANEGRNALHGGSHGLAVTRFNAV
    m144.pep
                 q144
                 AASYADNPFQINKQIGRVAGRIRGAAFDINGRTYRVEANEGRNALHGGSHGLAVTRFNAV
                        70
                                 80
                                          90
                                                  100
                                                           110
                       130
                                140
                                         150
                                                   160
                                                            170
                                                                     180
    m144.pep
                 AADGRSVVLRSRLATVGRRLSQRFGFGYFLPLGRGRPAYRYLSRHRARRHGVRPDAAHLL
                               ----GRRLSQRFG--YFLPLGRGRPAYRYLSRHRARRHGVRPDAAHLL
    q144
                                   130
                                              140
                                                       150
                                                                160
                       190
                                200
                                         210
    m144.pep
                AAGRGPARCGSAYSAGRTYAGRCRKTARLNGFRRPRSIX
                 g144
                AAGRGPARCGSAYSAGRTYSGRCRKTARLNGFRRPRSIX
                           180
                                     190
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 575>:
    al44.seg
              ATGAGCGATA CCCCCGCTAC CCGCGATTTC GGCCTGATCG ACGGGCGTGC
          51
              CGTAACCGGC TATGTGCTGT CCAACCGGCG TGGTACGCGT GTCTGCGTGC
              TGGACTTGGG CGGGATTGTG CAGGAATTTT CCGTTTTGGC AGACGGCGTG
         101
             CGCGAAAACC TCGTGGTGTC GTTCGACGAT GCGGCTTCCT ATGCGGACAA
         151
         201 TCCGTTTCAG ATTAACAAGC AGATAGGGCG CGTGGCCGGA CGCATCCGCG
         251 GTGCGGCGTT CGACATCAAC GGCAGGACTT ACCGCGTGGA GGCCAACGAA
             GGCAGGAACG CGCTGCACGG CGGTTCGCAC GGGCTGGCCG TTACCCGTTT
                                                                       - 3
         351
             CAACGCGGTG GCGCAGACG GCCGTTCGGT GGTGCTGCGC AGCCGCCTG.
         401 CAACAGTCGG CCGACGGTTA TCCCAACGAT TTGGATTTGG ATATTTCCTA
             CCGCTTGGAC GAGGACGACC GGCTTACCGT TACCTATCGC GCCACCGCGC
             TCGGCGACAC GGTGTTCGAC CCGACGCTGC ACATTTACTG GCGGCTGGAC
         551
             GCGGGCCTGC ACGATGCGGT TCTGCATATT CCGCAGGGCG GACATATTCC
              GGCCGATGCC GAAAAACTGC CCGTCTCAAC GGTTTCAGAC GACCTCGAAG
         601
         651
             TATTTGA
This corresponds to the amino acid sequence <SEQ ID 576; ORF 144.a>:
    al44.pep
             MSDTPATRDF GLIDGRAVTG YVLSNRRGTR VCVLDLGGIV QEFSVLADGV
           1
             RENLVVSFDD AASYADNPFQ INKQIGRVAG RIRGAAFDIN GRTYRVEANE
         101
             GRNALHGGSH GLAVTRFNAV AADGRSVVLR SRLXTVGRRL SQRFGFGYFL
             PLGRGRPAYR YLSRHRARRH GVRPDAAHLL AAGRGPARCG SAYSAGRTYS
         151
         201
             GRCRKTARLN GFRRPRSI*
                99.1% identity in 218 aa overlap
    ml44/al44
                        10
                                 20
                                          30
                                                   40
                {\tt MSDTPATRDFGLIDGRAVTGYVLSNRRGTRVCVLDLGGIVQEFSVLADGVRENLVVSFDD}
    m144.pep
                a144
                MSDTPATRDFGLIDGRAVTGYVLSNRRGTRVCVLDLGGIVQEFSVLADGVRENLVVSFDD
                       10
                                 20
                                          30
                                                   40
                                                            50
                                                                      60
                        70
                                 80
                                          90
                                                  100
                                                           110
                AASYADNPFQINKQIGRVAGRIRGAAFDINGRTYRVEANEGRNALHGGSHGLAVTRFNAV
    ml44.pep
```

BNSDOCID: <WO___9957280A2_I_>

a144	
m144.peg a144	130 140 150 160 170 180  AADGRSVVLRSRLATVGRRLSQRFGFGYFLPLGRGRPAYRYLSRHRARRHGVRPDAAHLL
4111	130 140 150 160 170 180
m144.pep	190 200 210 219  AAGRGPARCGSAYSAGRTYAGRCRKTARLNGFRRPRSIX
a144	AAGRGPARCGSAYSAGRTYSGRCRKTARLNGFRRPRSIX 190 200 210
The following	partial DNA sequence was identified in N. gonorrhoeae <seq 577="" id="">:</seq>
gl46.seq	partial Divisoquence was identified in N. gonorrhoeae <seq 577="" id="">:</seq>
9210.5cq 1	
51	
101	TACGUACIUS TACGUACIUS TURATURU COOMMO AANAA AAAA
151	TITE COLLEGE CELCOCKACE GEOCGE CECCCETTE COCCCCCCCCCCCCCCCCCCCCCCCCCC
201	
251	
301	
351	
401	
451	
501	TOTAL OCCUPACION DE LA CAMBANTA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTRA CONTR
551	
601	
This comes	
rins correspond	ls to the amino acid sequence <seq 146.ng="" 578;="" id="" orf="">:</seq>
g146.pep	_
1	MKQIPLRLLQ VVIDHDKVEQ YGLFDFMPCL RQPPLDNFPT VRPAPFEARG
51	MINERAL DIDSERORVA NI RRATINUDE ONUUTA CDDO DEUREDE
101	TAMILY CALO VOTER DVKKK I THE SUKKALMA VALCES EMBER'S BUSINESSES
151	VIIVOQQEARQ KGEARAGSGR NDKDVAFQIQ CHIEVIVIED DIVIGORDID
201	IFADAHILPL LF*
The following p	artial DNA sequence was identified in N. meningitidis <seq 579="" id="">:</seq>
m146.seq	
1	ATGGCGCAAA TCCTCCTCCG CTCGCGCCAA GTCGTCATTG ACCACGACAA
51	TACGGACTECATTECATE COMME CONTRACTOR
101	OFFICE CHICCOMMENT CONTROL CONTROL CONTROL CONTROL
151	THOTACGICG ARAGNAGGCG GCAGGATANA CATCCCCACC COMPAGGA
201	TOURS OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE P
251	TOTAL CONTROL OF CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTROL CONTR
301	TITLE TOCCA ANIACCTUGG CGTATTCCAA AAAACCTTCC TCCCCCTTC
351	COCCUMENCE CONTROL COCCUMENCE
401	TOOCTICAL GUGGUGGUG UGTUGTGTGC GACACCAAAA CCCCCAAAA
451	GITAIGGIII GCCAACAGCC GCGACACCAG CCTCCTTTTTTC CCCCTCCCCC
501	TAIGOGGCCGA AAIGAIAAAG ATGTTGCCTTT TTCCATAACC CCAGATTTC
551	TITACCIGIA INITITICCAG CCGATTGTAT CACAATCCAC ACCCACTOMORGO
601	CTATTTGCCG ATGCCCATAT TTTGCCGCTA TTGTTTTGA
	IGITIGA

### This corresponds to the amino acid sequence <SEQ ID 580; ORF 146>:

m146.pep 1 MAQILLRSRQ VVIDHDKVKQ YGLLDFMPCL RQPPLDNFPT VRPASVEARG
51 KYVERRRQDK DADGFGQRVA NLRRALNVDF QNHVIACRRQ RIHTLRACAV
101 IVAKYVGVFQ KSFLRDKRLK LFFGNKVIMY AVCFAFTRRA RRVRHGNAQT
151 VMVCQQPRHQ RGFARAGSGR NDKDVAFSIS GHIFYLYIFQ PIVSQWTPSF 101 151 201 LFADAHILPL LF*

Computer analysis of this amino acid sequence gave the following results: Homology with a predicted ORF from N. gonorrhoeae

m146 / g1	<b>46</b> 90.1% identity i	n 212	aa overlap			
	10	20	30	40	50	60
m146.pep	MAQILLRSRQVVIDHE	KVKQYO	LLDFMPCLRQP	PLDNFPTVRPA	SVEARGKYV	ERRRODK
g146	 MKQIPLRLLQVVIDHE	:   	:         :TEDEMPCTROD		:	
9140	10	20	30	40	SO 50	ERRRODK 60
m146.pep	70 DADGFGQRVANLRRAI	80 80	90 90 אינועדא כסססס דעי	100	110	120
m1 10.pcp	: :					
g146	DTDSFRQRVANLRRAI	NVDFQN	HVIACRRORIH	ALRACAVIVAE	YVCVFQKSL	LRDKRFK
	70	80	90	100	110	120
	130	140	150	160	170	180
m146.pep	LFFGNKVIMYAVCFAF	TRRARR	VRHGNAQTVMV	CQQPRHQRGFA	RAGSGRNDK	DVAFSIS
g146		 ם כמס כסידי	:			IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII
9110	130	140	150	160	170	180
	4.00					
m146.pep	190 GHIFYLYIFQPIVSQW	200 TPSELE	210			
mi io.pop						
g146	GHIFYLYIFQPIVSQR					
	190	200	210			
The following pa	artial DNA sequence v	vas ide	ntified in N.	meningitidis	<seo id<="" td=""><td>581&gt;:</td></seo>	581>:
al46.seq				_		
1	ATGGCGCAAA TCCTCCTC	CG CCC	GCGCCAA GTC	ATCATTG ACC	ACGACAA	
51 101	AATCGAACAA TACGGACT	GT TCG	ATTTCAT GCCT	TTGCCTT CGA	CAGCCTC	
151	CTTTGGATAA CTTCCCGA AAGCACATCG AAAGACGG	CG GCA	CGTCCCG CGTC	CCGTTGA GAC	GCGCAGC TCGGGCA	
201	GCGCATCTCG AACCTGAG	CC GCG	CCCTGAA CGT	CGATTTC CAA	AATCACG	<u>.</u> *.
251	TCATAACCTG CCGCCGCC	AA CGC	ATTCACA CCCT	CCGCGC TTG	TGCCGTA	***
301 351	ATAGTTGCCG AACACGTC	CG CGT	ATTCCAA AAAA	AGCCTCC TGC	GCGATAA	
401	GCGACTCAAA CTCTTCTT TCGCCTTCAC GCGGCGGA	TG GAA	ACAAAGT GATA	AATGTAC GCC	GTTTGCT	
451	GTTATGGTTT GCCAACAG	CC GCG	ACACCAG CGTC	GTTTTG CCC	GTGCCGG	
501	AAGCGGCCGA AATGATAA	AG ATG	TTGCCTT TTC	SATAAGC GGA	CATATTT	
551	TTTACCTGTA TATTTTCC	AG CCG	ATTGTAT CACA	AACGGAC ACC	CGGTTTC	
601	CTATTTGCCG ATGCCCAT	AT TTT	GCCGCTA TTGT	TTTGA		
This corresponds	s to the amino acid sec	mence	<seo 58<="" id="" td=""><td>2. ORF 146</td><td>a&gt;·</td><td></td></seo>	2. ORF 146	a>·	
a146.pep		1	52 4 12 50.	<b>2</b> , 01d 110.		
1	MAQILLRPRQ VIIDHDKI	EQ YGL	FDFMPCL RQPE	LDNFPT VRP	ASVETRS	
51	KHIERRRQDK DADGFGQR	IS NLS	RALNVDF QNHV	/ITCRRQ RIH	TLRACAV	
101	IVAEHVRVFQ KSLLRDKR	LK LFF	GNKVIMY AVC	FAFTRRT RRV	RHGNAQT	
151 201	VMVCQQPRHQ RGFARAGS	GR NDK	DVAFSIS GHIE	TYLYIFQ PIV	SQRTPGF	
201	LFADAHILPL LF*					
m146/a146	90.6% identity i	n 212	aa overlap			
	10	20	30	4.0	F.0	
m146.pep	MAQILLRSRQVVIDHD			40 PLDNFPTVRPA	50 SVEARGKYVI	60 ERRRODK
	111111	1::11	1:111111111		111:1:1:::	1111111
a146	<b>MAQILL</b> RPRQVIIDHD	KIEQYG	LFDFMPCLRQPE	PLDNFPTVRPA	SVETRSKHI	ERRRODK
	10	20	30	40	50	60
	70	80	90	100	110	120
		_				. 120

m146.pep	DADGFGQRVANLRRA	LNVDFQN	HVIACRRORIHTI	LRACAVIVA	KYVGVFQKSFL	RDKRLK
a146	::      DADGFGQRISNLSRAI	LNVDFQN: 80	:           HVITCRRQRIHTI 90	LRACAVIVAE	EHVRVFQKSLL	RDKRLK
	70	80	90	100	110	120
	130	140	150	160	170	180
m146.pep	LFFGNKVIMYAVCFA	TRRARR	VRHGNAQTVMVCC	QPRHQRGF#	RAGSGRNDKD	VAFSIS
a146	LFFGNKVIMYAVCFA	TRRTRR		 QPRHQRGFA	 RAGSGRNDKD	 VAFSIS
	130	140	150	160	170	180
	190	200	210			
m146.pep	GHIFYLYIFQPIVSQW		ADAHILPLLFX			
a146		:     TPGFLF7   200				

### The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 583>: g147.seq (partial)

```
..ATGCGACGAG AAGCCAAAAT GGCACAAATC ACACTCAAAC CCATTGTTTT
 1
       ATCAATTCTT TTAATCAACA CACCCCTCCT CGCCCAAGCG CATGAAACTG
 51
101
       AGCAATCGGT GGGCTTGGAA ACGGTCAGCG TCGTCGGCAA AAGCCGTCCG
151
       CGCGCGACTT CGGGGCTGCT GCACACTTCG ACCGCCTCCG ACAAAATCAT
201
       CTCCGGCGAT ACTTTGCGCC AAAAAGCCGT CAACTTGGGC GACGCTTTGG
      ACGGCGTACC GGGCATCCAC GCTTCGCAAT ACGGCGGCGG CGCATCCGCT
251
       CCCGTTATTC GCGGTCAAAC GGGCAGACGG ATTAAAGTAT TGAACCATCA
301
351
      CGGCGAAACG GGCGATATGG CGGACTTTTC TCCCGATCAC GCCATTATGG
401
      TAGATACCGC CTTGTCGCAA CAGGTTGAAA TCCTGCGCGG GCCGGTTACG
451
      CTCTTGTACA GCTCGGgcaa tgtggccgGG GCTGGtcaat gttgccgatg
      gAAAAAtccc ccaaaaAAtg cc..
```

#### This corresponds to the amino acid sequence <SEQ ID 584; ORF 147.ng>:

```
g147.pep (partial)

1 ..MRREAKMAQI TLKPIVLSIL LINTPLLAQA HETEQSVGLE TVSVVGKSRP
51 RATSGLLHTS TASDKIISGD TLRQKAVNLG DALDGVPGIH ASQYGGGASA
101 PVIRGQTGRR IKVLNHHGET GDMADFSPDH AIMVDTALSQ QVEILRGPVT
151 LLYSSGNVAG AGQCCRWKNP PKNA..
```

### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 585>: m147.seq (partial)

	\P				
1	CCGCATAAAA	CTGAGCAATC	GGTGGATTTG	GAAACGGTCA	GCGTCGTCGG
51	CAAAAGCCGT	CCGCGCGCCA			
101	CCGACAAAAT	CATCTCCGGC	GATACCTTGC	GCCAAAAAGC	CGTCAACTTG
151	GGCGACGCTT	TAGACGGCGT		CACGCTTCGC	
201	CGGCGCGTCT	GCTCCCGTCA		AACAGGCAGG	
251	TGTTGAACCA	TCACGGCGAA		TGGCGGATTT	TTCGCCCGAT
301	CACGCCATTA	TGGTAGATAC		CAACAGGTCG	
351	CGGGCCGGTT	ACGCTCTTGT		CAATGTGGCG	
401	ATGTTGCCGA	TGGCAAAATC			CGGCGTATCG
451	GGCGAACTCG	GATTGCGTTT			AACTCACGTC
501	CGGCGGCATC	AATATCGGTT		CTTTGTATTG	
551	GGCTGTACCG	CAAATCGGGG	GATTACGCCG		CCGCAATCTG
601	AAACGCCTGC	CCGACAGCCA		CAAACGGGCA	
651	GTCTTGGGTT	GGCGAAAAAG	_		AGCGACCGTC
701	GCGACCAATA	TGGTCTGCCT		ACGAATACGA	
751	GCCGACATCA	TCTGGCAAAA		AACAAACGCT	
801	TTATCCGCAC	CTGTTGACCG	AAGAAGACAT	CGATTACGAC	
851	TGAGCTGCGG	CTTCCACGAC	GACGATAATG		CACCCACAGC
901	GGCAGACCGT				
951	ATGGAAGCAA		GTTTTGAAGC		
1001	GCAACGACTA			CCTGCGCGTA	
1051	TTTAACAACC			GCGATGCAGT	CGAAAACTTT
			CCCCCCCATC	GWGTTGCGCC	ACCAACCCAT

1101	AGGTCGTCTG	AAAGGCAGCT	GGGGCGTGCA	ATATTTACAA	CAAAAATCCA
1151	GTGCTTTATC	TGCCATATCC	GAAGCGGTTA	AACAACCGAT	GCTGCTTGAC
1201	AACAAAGTGC	AACATTACAG	CTTTTTCGGT	GTAGAACAGG	CAAACTGGGA
1251	CAACTTCACG	CTTGAAGGAG	GCGTACGCGT	GGAAAAACAA	
1301	TTCAGTACGA	CAAAGCATTG	ATTGATCGGG	AAAACTACTA	CAACCACCCC
1351	CTGCCCGACC	TCGGCGCGCA	CCGCCAAACC	GCCCGCTCAT	TCGCACTTTC
1401	GGGCAACTGG	TATTTCACGC			ACCGCCTCCC
1451	ATCAGGAACG	CCTGCCGTCA	ACGCAAGAGC	TGTACGCACA	
1501	GTCGCCACCA		AGTCGGCAAC	AAACACCTCA	
1551	TTCCAACAAT	ATCGAACTCG	CGCTGGGCTA		CGCTGGCAAT
1601	ACAATCTGGC	ACTCTACCGC	AACCGCTTCG	GTAACTACAT	TTACGCCCAA
1651	ACCTTAAACG	ACGGACGCGG	CCCCAAATCC	ATCGAAGACG	ACAGCGAAAT
1701	GAAGCTCGTG	CGCTACAACC	AATCCGGCGC	CGACTTCTAC	GGCGCGGAAG
1751	GCGAAATCTA		ACACCGCGCT	ACCGCATCGG	CGTTTCCGGC
1801	GACTATGTAC	GAGGCCGTCT	GAAAAACCTG	CCTTCCCTAC	
1851	AGATGCCTAC	GGCAACCGTC	CTTTCATCGC	ACAGGACGAC	CCGGCAGAGA
1901	CCCGTGTTCC	GGCTGCGCGC	CTCGGCTTCC	ACCTGAAAGC	CAAAATGCCC
1951	GACCGTATCG	ATGCCAATTT	GGACTACTAC		CTCGCTGACC
2001	ACTCGCCCGC	TACGAAACGC		CGCGTGTTCG	CCCAAAACAA
2051	GCGCAAACTA	CCGCCGCAAT	GCACGCCCGG	ACACCATATG	CTCAACCTCG
2101	AAAGCCGACA	ACCTGCTCAA	ACGCGCTATG	GCGAGTGGAA	TTGGTACGTC
2151	CTCTGATACG	CCGCAAATGG	CCAATCCGTT		
2201	AGTTTTAA	CCGCMAATGG	GCCGCAGCTT	TACCGGCGGC	GTGAACGTGA
~~UI	VGITITW				

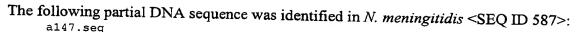
### This corresponds to the amino acid sequence <SEQ ID 586; ORF 147>: m147.pep (partial)

m147.pep	(partial)				
1	PHKTEQSVDL		PRATSGLLHT	STASDKIISG	DTLROKAVNL
51	GDALDGVPGI	HASQYGGGAS	APVIRGQTGR	RIKVLNHHGE	TGDMADFSPD
101	HAIMVDTALS	QQVEILRGPV	TLLYSSGNVA	GLVDVADGKI	PEKMPENGVS
151		NLEKLTSGGI	NIGLGKNFVL	HTEGLYRKSG	DYAVPRYRNL
201	KRLPDSHADS	QTGSIGLSWV	GEKGFIGVAY	SDRRDOYGLP	AHSHEYDDCH
251	ADIIWQKSLI	NKRYLQLYPH	LLTEEDIDYD	NPGLSCGFHD	DDNAHAHTHS
301	GRPWIDLRNK	RYELRAEWKQ	PFPGFEALRV	HLNRNDYRHD	EKAGDAVENE
351	FNNQTQNARI	ELRHQPIGRL	KGSWGVQYLQ	QKSSALSAIS	EAVKOPMLLD
401	NKVQHYSFFG	VEQANWDNFT	LEGGVRVEKQ	KASIOYDKAL	IDRENYYNHP
451	LPDLGAHRQT	ARSFALSGNW	YFTPQHKLSL	TASHQERLPS	TOELYAHGKH
501	VATNTFEVGN	KHLNKERSNN		RWQYNLALYR	
<b>5</b> 51	TLNDGRGPKS	IEDDSEMKLV		GAEGEIYFKP	
601	DYVRGRLKNL	PSLPGREDAY	GNRPFIAQDD		
651	DRIDANLDYY	RVFAQNKLAR	YETRTPGHHM	LNLGANYRRN	TRYGEWNWYV
701	KADNLLNQSV	YAHSSFLSDT	PQMGRSFTGG	VNVK <u>F</u> *	

# Computer analysis of this amino acid sequence gave the following results: Homology with a predicted ORF from N. gonorrhoeae

m147 / g147 92.3% identity in 142 aa overlap

4.5				10	20	30
m147.pep			PHKT	<b>EQSVDLETVS</b>	VVGKSRPRAT	SGLLHTS
				1111 11111	1111111111	111111
g147	MRREAKMAQITLK	PIVLSILLIN	TPLLAQAHET:	<b>EQSVGLETVS</b>	VVGKSRPRAT	SGLLHTS
	10	20	30	40	50	60
	40	50	60	70	80	90
m147.pep	TASDKIISGDTLR	QKAVNLGDAL	DGVPGIHASO	YGGGASAPVII	RGOTGRRIKV	LNHHGET
	11111111111	11111111111	пппппі	111111111		
g147	TASDKIISGDTLR	QKAVNLGDAL	DGVPGIHASO	YGGGASAPVI	RGOTGRRIKV	LNHHGET
	70	80	90	100	110	120
	100	110	120	130	140	150
m147.pep	GDMADFSPDHAIM	VDTALSQQVE	ILRGPVTLLY	SSGNVAGLVDV	JADGKI PEKM	PENGVSG
1.2		1	171411711		1 1 1	
g147	GDMADFSPDHAIM	VDTALSQQVE	ILRGPVTLLY	SSGNVAGAGO	CCRWKNPPKN	A



ATGCGACGAG AAGCCAAAAT GGCACAAACT ACACTCAAAC CCATTGTTTT ATCAATTCTT TTAATCAACA CACCCCTCCT CTCCCAAGCG CATGGAACTG 51 101 AGCAATCAGT GGGCTTGGAA ACGGTCAGCG TCGTCGGCAA AAGCCGTCCG CGCGCCACTT CGGGGCTGCT GCACACTTCT ACCGCCTCCG ACAAAATCAT CAGCGGCGAC ACCTTGCGAC AAAAAGCCGT CAACTTGGGT GATGCTTTAG 201 ACGGCGTACC GGGCATTCAT GCCTCGCAAT ACGGCGGCGG CGCATCCGCT CCCGTTATTC GCGGTCAAAC AGGCAGACGG ATTAAAGTGT TGAACCATCA 301 CGGCGAAACG GGCGACATGG CGGACTTCTC TCCAGACCAT GCAATCATGG 351 TGGACAGCGC CTTGTCGCAA CAGGTCGAAA TCCTGCGCGG TCCGGTTACG 401 CTCTTGTACA GCTCGGGCAA TGTGGCGGGG CTGGTCGATG TTGCCGATGG 451 CAAAATCCCC GAAAAAATGC CTGAAAACGG CGTATCGGGC GAACTCGGAT 501 TGCGTTTGAG CAGCGGCAAT CTGGAAAAAC TCACGTCCGG CGGCATCAAT 551 ATCGGTTTGG GCAAAAACTT TGTATTGCAC ACGGAAGGGC TGTACCGCAA 601 ATCGGGGGAT TACGCCGTAC CGCGTTACCG CAATCTGAAA CGCCTGCCCG 651 701 ACAGCCACGC CGATTCGCAA ACGGGCAGCA TCGGGCTGTC TTGGGTTGGC GAAAAAGGCT TTATCGGCGC AGCATACAGC GACCGTCGCG ACCAATATGG TCTGCCTGCC CACAGCCACG AATACGATGA TTGCCACGCC GACATCATCT GGCAAAAGAG TTTGATTAAC AAACGCTATT TGCAGCTTTA TCCGCACCTG 851 TTGACCGAAG AAGACATCGA TTACGACAAT CCGGGCTTGA GCTGCGGCTT TCACGACGAC GATGATGCAC ACGCCCATGC CCACAACGGC AAACCTTGGA 951 TAGACCTGCG CAACAAACGC TACGAACTCC GCGCCGAATG GAAGCAACCG 1001 TTCCCCGGTT TTGAAGCCCT GCGCGTACAC CTGAACCGCA ACGACTACCG 1051 CCACGACGAA AAAGCAGGCG ATGCAGTAGA AAACTTTTTT AACAACCAAA 1101 CGCAAAACGC CCGTATCGAG TTGCGCCACC AACCCATAGG CCGTCTGAAA 1151 GGCAGCTGGG GCGTGCAATA TTTGGGACAA AAATCCAGTG CTTTATCTGC 1201 CACATCCGAA GCGGTCAAAC AACCGATGCT GCTTGACAAT AAAGTGCAAC ATTACAGCTT TTTCGGTGTA GAACAGGCAA ACTGGGACAA CTTCACGCTT 1301 GAAGGCGGCG TACGCGTGGA AAAACAAAAA GCCTCCATCC GCTACGACAA 1351 AGCATTGATT GATCGGGAAA ACTACTACAA CCATCCCCTG CCCGACCTCG 1401 GCGCGCACCG CCAAACCGCC CGCTCATTCG CACTTTCGGG CAACTGGTAT 1451 1501 TTCACGCCAC AACACAAACT CAGCCTGACC GCCTCCCATC AGGAACGCCT 1551 GCCGTCAACG CAAGAGCTGT ACGCACACG CAAACACGTC GCCACCAACA CCTTTGAAGT CGGCAACAAA CACCTCAACA AAGAGCGTTC CAACAATATC 1601 GAACTCGCGC TGGGCTACGA AGGCGACCGC TGGCAATACA ATCTGGCACT 1651 CTACCGCAAC CGCTTCGGCA ACTACATTTA CGCCCAAACC TTAAACGACG 1701 GACGCGGCCC CAAATCCATC GAAGACGACA GCGAAATGAA GCTCGTGCGC TACAACCAAT CCGGTGCGGA CTTCTACGGC GCGGAAGGCG AAATCTACTT 1801 CAAACCGACA CCGCGCTACC GCATCGGCGT TTCCGGCGAC TATGTACGAG 1901 GCCGTCTGAA AAACCTGCCT TCCCTACCCG GCAGGGAAGA CGCCTACGGC 1951 AACCGCCCAC TCATTGCCCA AGCCGACCAA AACGCCCCTC GCGTTCCGGC TGCGCGCCTC GGCGTCCACC TGAAAGCCTC GCTGACCGAC CGCATCGATG 2001 CCAATTTGGA CTACTACCGC GTGTTCGCCC AAAACAAACT CGCCCGCTAC 2051 2101 GAAACGCGCA CGCCCGGACA CCATATGCTC AACCTCGGCG CAAACTACCG CCGCAATACG CGCTATGGCG AGTGGAATTG GTACGTCAAA GCCGACAACC 2151 TGCTCAACCA ATCCGTTTAC GCCCACAGCA GCTTCCTCTC TGATACGCCG 2201 CAAATGGGCC GCAGCTTTAC CGGCGGCGTG AACGTGAAGT TTTAA

### This corresponds to the amino acid sequence <SEQ ID 588; ORF 147.a>:

/.pep					
1	MRREAKMAQT		LINTPLLSOA	HGTEOSVGLE	TVSVVGKSRP
51	RATSGLLHTS	TASDKIISGD	TLROKAVNIG	DALDGVPGTH	A COVCCCA CA
101	FAIRGOIGER	TKATNHHGEL	GDMADFSPDH	ATMVDSALSO	OWETTOCOUM
151	TTISSGNAM	LVDVADGKIP	EKMPENGVSG	ELGLELSSON	TENT TOCCOL
201	TOTOVNEATH	TEGLYRKSGD	YAVPRYRNIK	RI.PDSHADSO	TOSTOT CHICA
251	ENGRIGARIS	DRRDQYGLPA	HSHEYDDCHA	DITWOKSTIM	PRVI OF VIII
301	PIEEDIDID	PGLSCGFHDD	DDAHAHAHNG	KPWIDI.DNKD	VEIDABWEOD
351	FEGEFALKAN	LNKNDYRHDE	KAGDAVENFF	NNOTONARIE	LRHOPIGRLK
401	GOMGAĞITGĞ	KSSALSATSE	AVKOPMLLDN	KVQHYSFFGV	ECANWONETT
451	EGGVRVEKQK	ASIRYDKALI	DRENYYNHPI.	PDLGAHRQTA	RSFALSGNWY
501	FTPQHKLSLT	ASHQERLPST	QELYAHGKHV	ATMTERUCKE	HT MICED CANA
551	ELALGYEGDR	WQYNLALYRN	RFGNYIYAQT	LNDGRGPKSI	EDDSEMKTAR

601 651 701 751	YNQSGADFYG AEGEIYFKPT PRYRIGVSGD YVRGRLKNLP SLPGREDAYG NRPLIAQADQ NAPRVPAARL GVHLKASLTD RIDANLDYYR VFAQNKLARY ETRTPGHHML NLGANYRRNT RYGEWNWYVK ADNLLNQSVY AHSSFLSDTP QMGRSFTGGV NVK <u>F</u> *
m147/a147	98.1% identity in 734 aa overlap
m147.pep	10 20 30 PHKTEQSVDLETVSVVGKSRPRATSGLLHTS
	10 20 30 40 50 60 40. 50 60 70 80 90
m147.pep	40 50 60 70 80 90 TASDKIISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGET
a147	TASDKIISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGET 70 80 90 100 110 120
m147.pep	100 110 120 130 140 150 GDMADFSPDHAIMVDTALSQQVEILRGPVTLLYSSGNVAGLVDVADGKIPEKMPENGVSG
a147	
m147.pep	160 170 180 190 200 210 ELGLRLSSGNLEKLTSGGINIGLGKNFVLHTEGLYRKSGDYAVPRYRNLKRLPDSHADSQ
a147	
m147.pep	220 230 240 250 260 270 TGSIGLSWVGEKGFIGVAYSDRRDQYGLPAHSHEYDDCHADIIWQKSLINKRYLQLYPHL
a147	
m147.pep	280 290 300 310 320 330 LTEEDIDYDNPGLSCGFHDDDNAHAHTHSGRPWIDLRNKRYELRAEWKQPFPGFEALRVH
a147	
m147.pep	340 350 360 370 380 390 LNRNDYRHDEKAGDAVENFFNNQTQNARIELRHQPIGRLKGSWGVQYLQQKSSALSAISE
a147	
m147.pep	400 410 420 430 440 450 AVKQPMLLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASIQYDKALIDRENYYNHPL
a147	
m147.pep	460 470 480 490 500 510 PDLGAHRQTARSFALSGNWYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNK
a147	
	490     500     510     520     530     540       520     530     540     550     560     570
m147.pep	HLNKERSNNIELALGYEGDRWQYNLALYRNRFGNYIYAQTLNDGRGPKSIEDDSEMKLVR
a147	HLNKERSNNIELALGYEGDRWQYNLALYRNRFGNYIYAQTLNDGRGPKSIEDDSEMKLVR

•	550	560	570	580	590	<b>60</b> 0
m147.pep	580	590	600	610	620	630
pcp	YNOSGADFYGAEG	CITICKPTPRY	RIGVSGDYVRO	RLKNLPSLP	GREDAYGNRP:	FIAQDDQ
a147	VNOSCA DEVCA PO	]	<u> </u>			:111 11
411	YNQSGADFYGAEGI	STALKELEKA	RIGVSGDYVRO	RLKNLPSLP	GREDAYGNRP:	LIAQADQ
	610	620	630	640	650	660
	640	650	660	670	680	690
m147.pep	NAPRVPAARLGFHI	LKASLTDRID	ANLDYYRVFAC	NKLARYETR	PECHHMINIG	MYDDNT
			1111111111	11111111		111111
a147	NAPRVPAARLGVHI	KASLTDRID	ANLDYYRVFAC	NKLARYETRI	PGHHMI.NI.GI	TIIIIII
	670	680	690	700	710	720
	700	710	720	730		
m147.pep	RYGEWNWYVKADNI	LNQSVYAHS	SFLSDTPQMGR	SFTGGVNVK	rx:	
				111111111	1	
a147	RYGEWNWYVKADNI	LNQSVYAHS	SFLSDTPOMGR	SFTGGVNVK	י ציי	
	730	740	750	760	41	

#### The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 589>:

```
1 ATGGCGTTAA AAACATCAAA CTTGGAACAC GCAATGCtgg ttcaTCCCGA
51 AgctATgagt gtcggegCGC TTGccgAcaa AATCCGCAAA AtcgaAAact
101 gGCCGCAAAA AGgCaTCTTA TTCCACGACA TCACGCCCGT CCTGCAAAGT
151 GCGGAATACT TCCGCCTTTT GGTCGATTTG CTGGTTTACC GCTATATGGA
201 TCAGAAAATC GACATCGTTG CCGGCTTGGA CGCGCGCGC TTCATTATCG
251 GCGCGCACT CGCCTACCAG CTCAaCGtcg gctTCGTCC CATCCGCAAA
301 AAAGGCAAGC TGCCTTTTGA AACCGTATCG CAAAGCTACG CGCGCGAA
351 CGGGGAAGCT GCGGTGGAAA TCCACACCGa tgccgTCAAA CCCGGTTCGC
401 GCGTCCTGCT GGTCGATGAT TTGGTTGCCA CGGCGGCAC AATGCTTGCC
451 GGGCTGGAAC TGATCCGCAA ACTCGGCGGG GAAATTGTCG AAGCCTGCC
551 GCGCGCCCTT ATTTACCCTG CTTCAAAACG AAGGCTGCAT GAAAGGCTGA
```

### This corresponds to the amino acid sequence <SEQ ID 590; ORF 148.ng>: g148.pep

1 MALKTSNLEH AMLVHPEAMS VGALADKIRK IENWPQKGIL FHDITPVLQS 51 AEYFRLLVDL LVYRYMDQKI DIVAGLDARG FIIGAALAYQ LNVGFVPIRK 101 KGKLPFETVS QSYALEYGEA AVEIHTDAVK PGSRVLLVDD LVATGGTMLA 151 GLELIRKLGG EIVEAAAILE FTDLQGGKNI RASGAPLFTL LQNEGCMKG*

### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 591>:

ATGGCGTTAA AAACATCAAA CTTGGAACAC GCAATGCTGG TTCATCCCGA 1 AGCTATGAGT GTCGGCGCGC TTGCCGACAA AATCCGCAAA ATCGAAAACT 101 GGCCGCAAAA AGGCATCTTA TTCCACGACA TCACGCCCGT CCTTCAAAGC 151 GCGGAATACT TCCGCCTTTT GGTTGATTTA TTGGTTTACC GCTATATGGA 201 TCAGAAAATC GACATCGTTG CCGGTTTGGA CGCGCGCGC TTCATTATCG 251 GCGCGGCACT CGCCTACCAG CTCAACGTCG GTTTCGTCCC CATCCGCAAA 301 AAAGGCAAGC TGCCTTTTGA AACCGTATCG CAAAGCTACG CGCTCGAATA 351 CGGGGAAGCT GCGGTGGAAA TCCACACCGA TGCCGTCAAA CTCGGTTCGC GCGTGCTGCT GGTCGATGAT TTGATTGCCA CGGGCGGCAC GATGCTTGCC 401 451 GGACTGGAAC TGATCCGCAA ACTCGGCGGA GAAATTGTCG AAGCCGCCGC 501 CATTTTGGAA TTTACCGACC TTCAAGGCGG CAAGAATATC CGTGCAAGCG 551 GCGCGCCCTT ATTTACCCTG CTTCAAAACG AAGGCTGTAT GAAGGGCTGA

#### This corresponds to the amino acid sequence <SEQ ID 592; ORF 148>:

1 MALKTSNLEH AMLVHPEAMS VGALADKIRK IENWPQKGIL FHDITPVLQS 51 AEYFRLLVDL LVYRYMDQKI D<u>IVAGLDARG FIIGAALA</u>YQ LNVGFVPIRK KGKLPFETVS QSYALEYGEA AVEIHTDAVK LGSRVLLVDD LIATGGTMLA GLELIRKLGG EIVEAAAILE FTDLQGGKNI RASGAPLFTL LQNEGCMKG*

Computer analysis of this amino acid sequence gave the following results: Homology with a predicted ORF from N. gonorrhoeae

m148 / g1	<b>48</b> 99.0% identity in	199 aa over	lap		
	10	20 3	0 40	50	60
m148.pep	MALKTSNLEHAMLVHPE				
~1.40	MALKECAL FURNI VUDE				
g1 <b>4</b> 8	MALKTSNLEHAMLVHPE 10		RIENWPOKGIL.	FHDITPVLQSAF 50	YFRLLVDL 60
				30	00
_	70		100	110	120
m148.pep	LVYRYMDQKIDIVAGL	ARGFIIGAALAY	QLNVGFVPIRK	KGKLPFETVSQS	YALEYGEA
g148		IIIIIIIIIIIIIIIIIIIIIIIIIIIAAIAY	OINVGEVPTRKI		YALEYGEA
9	70		0 100	110	120
			_		
m148.pep	130 AVEIHTDAVKLGSRVLI	140 15		170	180
m140.pep		.vDDLIAIGGIML	HGLELIKKLGG	FIARAATTEET	DLQGGKNI
g148	AVEIHTDAVKPGSRVLI	VDDLVATGGTML	AGLELIRKLGG	EIVEAAAILEFT	DLQGGKNI
	130	140 15	0 160	170	180
	190	200			
m148.pep	RASGAPLFTLLONEGCN				
• -					
g148	RASGAPLFTLLQNEGCN				•
	190	200			
The following n	artial DNA sequence w	as identified i	n <i>N meninoi</i>	tidis <seo i<="" td=""><td>D 593&gt;·</td></seo>	D 593>·
a148.seq				52	D 333°.
i	ATGGCGTTAA AAACATCAA				
51	AGCTATGAGT GTCGGTGCG	C TTGCCGACAA	AATCCGCAAA	ATCGAAAACT	
101 151	GGCCGCAAAA AGGCATCTT GCGGAATACT TCCGACTTT				ě.
201	TCAGAAAATC GACATCGTT				
251	GCGCGGCACT CGCCTACCA	G CTCAACGTCG	GTTTCGTCCC	CATCCGCAAA	
301	AAAGGCAAGC TGCCTTTTG				
351	CGGGGAAGCT GCGGTGGAA				
401 451	GCGTGCTGCT GGTCGATGA GGACTGGAGC TGATCCGCA				
501	CATTTTGGAA TTTACCGAC	C TTCAAGGCGG	CAAGAATATC	CGTGCAAGCG	
551	GCGCGCCCTT ATTTACCCT				
This correspond	s to the amino acid seq	uence <seo i<="" th=""><th>D 504: OPF</th><th>148 2&gt;</th><th></th></seo>	D 504: OPF	148 2>	
a148.pep	s to the armie acid seq	defice -BLQ i	D 334, OKI	170.4	
1	MALKTSNLEH AMLVHPEAN	S VGALADKIRK	IENWPOKGIL	FHDITPVLOS	
51	AEYFRLLVDL LVYRYMDQ	I D <u>IVAGLDAR</u>	FIIGAALAYQ	LNVGFVPIRK	
	KGKLPFETVS QSYALEYGE	A AVEIHTDAVK	LGSRVLLVDD	LVATGGTMLA	
151	GLELIRKLGG EIVEAAAII	E FTDLQGGKNI	RASGAPLFTL	LQNEGCMKG*	
m148/a148	99.5% identity in	199 aa over	lap		
	10	20 3	0 40	50	60
m148.pep	MALKTSNLEHAMLVHPE	CAMSVGALADKIR	KIENWPQKGIL	FHDITPVLQSAE	EYFRLLVDL
0140	MAI KMCNI EUANI VUDE	111111111		11111111111	1111111
a148	MALKTSNLEHAMLVHPE 10		KIENWPQKGIL:	FHDITPVLQSAE 50	CYFRLLVDL 60
				30	00

70

80

90

100

110

120

WO 99/57280



```
m148.pep
          {\tt LVYRYMDQKIDIVAGLDARGFIIGAALAYQLNVGFVPIRKKGKLPFETVSQSYALEYGEA}
           a148
          LVYRYMDQKIDIVAGLDARGFIIGAALAYQLNVGFVPIRKKGKLPFETVSQSYALEYGEA
                 70
                         80
                                 90
                                        100
                                                        120
                130
                        140
                                150
                                        160
                                                170
          AVEIHTDAVKLGSRVLLVDDLIATGGTMLAGLELIRKLGGEIVEAAAILEFTDLQGGKNI
m148.pep
          AVEIHTDAVKLGSRVLLVDDLVATGGTMLAGLELIRKLGGEIVEAAAILEFTDLQGGKNI
a148
                        140
                                150
                                        160
                190
m148.pep
          RASGAPLFTLLQNEGCMKGX
          1111111111111111111111
a148
          RASGAPLFTLLQNEGCMKGX
                190
```

The following partial DNA sequence was identified in N.gonorrhoeae <SEQ ID 595>:

```
ATGTTGATTG ACAACAATGT CCGCCATTAC AGCTTTTTCG GTGTAGAACA
   1
  51 GGCAAATTGG GACAACTTCA CGCTTGAAGG CGGCGTACGC GTGGAAAAAC
 101 AAAAAGCCTC CATCCGGTAC GACAAAGCAT TGATTGATCG AGAAAACTAC
 151 TACAACCAGC CCCTGCCCGA CCTCGGCGCG CACCGCCAAA CCGCCCGCTC
     GTTCGCACTT TCGGGCAACT GGTATTTCAC GCCACACCAC AAACTCAGCC
      TGACCGCCTC CCATCAGGAa cgCCTGCCGT CAACGCaagA actGtACgca
 251
      cacggcAAGC ACGtcgccac CAACACCTTT GAagtcggca acaaACACCT
 301
 351 CAACAAAGAG CgttccaacA atatcgaACT CGCGCTGGgc tAcaaaggcg
 401 accGCTGGCA ATACAATCTG GCAGCCTACC GCAACCGAET CGGCAACTAC
 451 ATTTACGCCC AAACCTTAaa cgacggacgC GGCCCCAAAT CCATCgaaga
 501 CGacagcgaA ATGaagcTCG TGCGCTACAA CCAATCCGGT GCCGACTTCT
 551 ACGGCGCGA aggCGaaatc tACTTCaaaC CGACACCGCG CTACCGCATC
 601 GGTGTTTCCG GCGACTatgt acgaggccgT CTGAAAAACC TGCCGTCCCT
     ACCCGGCAGG gaagatccCT AcggcAAACG TCccttcaTC GCACAAGCCG
 651
     ACCAAAACGC CCCCCGCATT ccggctGCGC GCCTCGGCTT CCACCTGAAA
 751
     ACCTCGCTAA CCGACCGTAT CGATGCCAAT TTGGACTACT ACCGCGTGTT
801 CGCCCAAAAC AAACTCGCCC GCTACGAAAC GCGTACGCCC GGACACCATA
851 TGCTCAACCT CGGTGCAAAC TACCGCCGCA ATACGCGCTA TGGCGAGTGG
901 AATTGGTACG TCAAAGCCGA CAACCTGCtc aACcaatCcg tTTACGCCCa
951 CAGCAGCTTC CTCTCTGATA CGCCGCAAAt gGGCCGCAGC TTtgccgGCg
1001 gcgtaAACGT GaAGTTttaA
```

### This corresponds to the amino acid sequence <SEQ ID 596; ORF 149.ng>:

```
MLIDNNVRHY SFFGVEQANW DNFTLEGGVR VEKQKASIRY DKALIDRENY
YNOPLPDLGA HRQTARSFAL SGNWYFTPHH KLSLTASHQE RLPSTQELYA
HGKHVATNTF EVGNKHLNKE RSNNIELALG YKGDRWQYNL AAYRNRFGNY
IYAQTLNDGR GPKSIEDDSE MKLVRYNQSG ADFYGAEGEI YFKPTPRYRI
GVSGDYVRGR LKNLPSLPGR EDPYGKRPFI AQADQNAPRI PAARLGFHLK
TSLTDRIDAN LDYYRVFAQN KLARYETRTP GHHMLNLGAN YRRNTRYGEW
NWYVKADNLL NQSVYAHSSF LSDTPQMGRS FAGGVNVKF*
```

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 597>:

```
1 ATGCTGCTTG ACAACAAGT GCAACATTAC AGCTTTTCG GTGTAGAACA
51 GGCAAACTGG GACAACTTCA CGCTTGAAGG AGGCGTACGC GTGGAAAAAC
101 AAAAAGCCTC CATTCAGTAC GACAAAGCAT TGATTGATCG GGAAAACTAC
151 TACAACCACC CCCTGCCCGA CCTCGGCGCG CACCGCCAAA CCGCCCGCTC
201 ATTCGCACTT TCGGGCAAACT GGTATTTCAC GCCACAACAC AAACTCAGCC
251 TGACCGCCTC CCATCAGGAA CGCCTGCCGT CAACGCAAGA GCTGTACGCA
301 CACGGCAACA ACGTCGCCAC CAACACCTTT GAAGTCGGCA ACAAACACCT
351 CAACAAAGAG CGTTCCAACA ATATCGAACT CGCGCTGGCC TACGAAGGCG
401 ACCGCTGGCA ATACAATCTG GCACTCTACC GCAACCGCTT CGGTAACTAC
```

```
ATTTACGCCC AAACCTTAAA CGACGGACGC GGCCCCAAAT CCATCGAAGA
         CGACAGCGAA ATGAAGCTCG TGCGCTACAA CCAATCCGGC GCCGACTTCT
     551
         ACGGCGCGGA AGGCGAAATC TACTTCAAAC CGACACCGCG CTACCGCATC
         GGCGTTTCCG GCGACTATGT ACGAGGCCGT CTGAAAAACC TGCCTTCCCT
         ACCCGGCAGA GAAGATGCCT ACGGCAACCG TCCTTTCATC GCACAGGACG
         ACCAAAATGC CCCCCGTGTT CCGGCTGCGC GCCTCGGCTT CCACCTGAAA
         GCCTCGCTGA CCGACCGTAT CGATGCCAAT TTGGACTACT ACCGCGTGTT
     801
         CGCCCAAAAC AAACTCGCCC GCTACGAAAC GCGCACGCCC GGACACCATA
     851
         TGCTCAACCT CGGCGCAAAC TACCGCCGCA ATACGCGCTA TGGCGAGTGG
         AATTGGTACG TCAAAGCCGA CAACCTGCTC AACCAATCCG TTTACGCCCA
     901
         CAGCAGCTTT CTCTCTGATA CGCCGCAAAT GGGCCGCAGC TTTACCGGCG
     951
         GCGTGAACGT GAAGTTTTAA
This corresponds to the amino acid sequence <SEQ ID 598; ORF 149>:
m149.pep
         MLLDNKVQHY SFFGVEQANW DNFTLEGGVR VEKQKASIQY DKALIDRENY
      1
         YNHPLPDLGA HRQTARSFAL SGNWYFTPQH KLSLTASHQE RLPSTQELYA
     51
         HGKHVATNTF EVGNKHLNKE RSNNIELALG YEGDRWQYNL ALYRNRFGNY
         IYAQTLNDGR GPKSIEDDSE MKLVRYNQSG ADFYGAEGEI YFKPTPRYRI
    151
    201 GVSGDYVRGR LKNLPSLPGR EDAYGNRPFI AQDDQNAPRV PAARLGFHLK
    251 ASLTDRIDAN LDYYRVFAQN KLARYETRTP GHHMLNLGAN YRRNTRYGEW
    301 NWYVKADNLL NQSVYAHSSF LSDTPQMGRS FTGGVNVKF*
Computer analysis of this amino acid sequence gave the following results:
Homology with a predicted ORF from N.gonorrhoeae
ORF 149 shows 95.9% identity over a 339 aa overlap with a predicted ORF (ORF 149.ng)
from N. gonorrhoeae:
m149/g149
                   10
                             20
                                      30
                                               40
                                                        50
                                                                 60
            MLLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASIQYDKALIDRENYYNHPLPDLGA
m149.pep
            MLIDNNVRHYSFFGVEQANWDNFTLEGGVRVEKQKASIRYDKALIDRENYYNQPLPDLGA
g149
                   10
                            20
                                      30
                                               40
                                                        50
                                                                 60
                   70
                             80
                                      90
                                              100
                                                       110
                                                                120
           HRQTARSFALSGNWYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKE
m149.pep
                                                                         4
            HRQTARSFALSGNWYFTPHHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKE
g149
                   70
                            80
                                      90
                                              100
                                                       110
                                                                120
                  130
                            140
                                     150
                                              160
                                                       170
                                                                180
m149.pep
            RSNNIELALGYEGDRWQYNLALYRNRFGNYIYAQTLNDGRGPKSIEDDSEMKLVRYNOSG
            g149
            RSNNIELALGYKGDRWQYNLAAYRNRFGNYIYAQTLNDGRGPKSIEDDSEMKLVRYNQSG
                  130
                           140
                                     150
                                              160
                                                       170
                                                                180
                           200
                                     210
                                              220
                                                       230
                                                                240
m149.pep
           ADFYGAEGEIYFKPTPRYRIGVSGDYVRGRLKNLPSLPGREDAYGNRPFIAODDONAPRV
            g149
           ADFYGAEGEIYFKPTPRYRIGVSGDYVRGRLKNLPSLPGREDPYGKRPFIAQADQNAPRI
```

190

250

250

310

200

260

260

320

NWYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKFX

210

270

270

330

PAARLGFHLKASLTDRIDANLDYYRVFAQNKLARYETRTPGHHMLNLGANYRRNTRYGEW

PAARLGFHLKTSLTDRIDANLDYYRVFAQNKLARYETRTPGHHMLNLGANYRRNTRYGEW

220

280

280

230

290

240

300

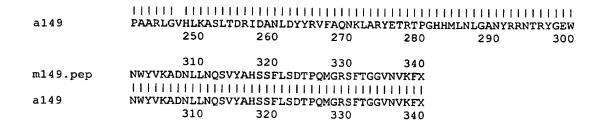
BNSDOCID: <WO___9957280A2_I_>

m149.pep

m149.pep

g149

g149 NWYVKADNLLNQSVYAHSSFLSDTPQMGRSFAGGVNVKFX 310 320 330 The following partial DNA sequence was identified in N. meningitidis <SEQ ID 599>: al49.seg ATGCTGCTTG ACAATAAAGT GCAACATTAC AGCTTTTTCG GTGTAGAACA 1 GGCAAACTGG GACAACTTCA CGCTTGAAGG CGGCGTACGC GTGGAAAAAC 51 AAAAAGCCTC CATCCGCTAC GACAAAGCAT TGATTGATCG GGAAAACTAC 101 TACAACCATC CCCTGCCCGA CCTCGGCGCG CACCGCCAAA CCGCCCGCTC 151 ATTCGCACTT TCGGGCAACT GGTATTTCAC GCCACAACAC AAACTCAGCC 201 TGACCGCCTC CCATCAGGAA CGCCTGCCGT CAACGCAAGA GCTGTACGCA 251 CACGGCAAAC ACGTCGCCAC CAACACCTTT GAAGTCGGCA ACAAACACCT 301 CAACAAGAG CGTTCCAACA ATATCGAACT CGCGCTGGGC TACGAAGGCG 351 ACCGCTGGCA ATACAATCTG GCACTCTACC GCAACCGCTT CGGCAACTAC 401 ATTTACGCCC AAACCTTAAA CGACGGACGC GGCCCCAAAT CCATCGAAGA 451 CGACAGCGAA ATGAAGCTCG TGCGCTACAA CCAATCCGGT GCGGACTTCT 501 551 ACGGCGCGGA AGGCGAAATC TACTTCAAAC CGACACCGCG CTACCGCATC GGCGTTTCCG GCGACTATGT ACGAGGCCGT CTGAAAAACC TGCCTTCCCT 601 ACCCGGCAGG GAAGACGCCT ACGGCAACCG CCCACTCATT GCCCAAGCCG 651 ACCAAAACGC CCCTCGCGTT CCGGCTGCGC GCCTCGGCGT CCACCTGAAA 701 751 GCCTCGCTGA CCGACCGCAT CGATGCCAAT TTGGACTACT ACCGCGTGTT CGCCCAAAAC AAACTCGCCC GCTACGAAAC GCGCACGCCC GGACACCATA 851 TGCTCAACCT CGGCGCAAAC TACCGCCGCA ATACGCGCTA TGGCGAGTGG 901 AATTGGTACG TCAAAGCCGA CAACCTGCTC AACCAATCCG TTTACGCCCA CAGCAGCTTC CTCTCTGATA CGCCGCAAAT GGGCCGCAGC TTTACCGGCG 1001 GCGTGAACGT GAAGTTTTAA This corresponds to the amino acid sequence <SEQ ID 600; ORF 149.a>: a149.pep MLLDNKVQHY SFFGVEQANW DNFTLEGGVR VEKQKASIRY DKALIDRENY YNHPLPDLGA HRQTARSFAL SGNWYFTPQH KLSLTASHQE RLPSTQELYA 51 HGKHVATNTF EVGNKHLNKE RSNNIELALG YEGDRWQYNL ALYRNRFGNY 101 IYAQTLNDGR GPKSIEDDSE MKLVRYNQSG ADFYGAEGEI YFKPTPRYRI 201 GVSGDYVRGR LKNLPSLPGR EDAYGNRPLI AQADQNAPRV PAARLGVHLK ASLTDRIDAN LDYYRVFAQN KLARYETRTP GHHMLNLGAN YRRNTRYGEW NWYVKADNLL NQSVYAHSSF LSDTPQMGRS FTGGVNVK $\underline{\mathbf{F}}^{\star}$ m149/a149 98.8% identity in 339 aa overlap 10 20 30 40 MLLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASIQYDKALIDRENYYNHPLPDLGA m149.pep MLLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASIRYDKALIDRENYYNHPLPDLGA a149 10 20 30 40 50 70 80 90 100 HRQTARSFALSGNWYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKE 110 m149.pep allandida kalanda kalanda kalanda kalanda kalanda kalanda kalanda kalanda kalanda kalanda kalanda kalanda kala HRQTARSFALSGNWYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKE a149 70 80 90 100 120 130 140 150 160 170 RSNNIELALGYEGDRWQYNLALYRNRFGNYIYAQTLNDGRGPKSIEDDSEMKLVRYNQSG m149.pep  $oldsymbol{n}$ RSNNIELALGYEGDRWQYNLALYRNRFGNYIYAQTLNDGRGPKSIEDDSEMKLVRYNQSG a149 130 140 150 170 180 190 200 210 220 230 ADFYGAEGEIYFKPTPRYRIGVSGDYVRGRLKNLPSLPGREDAYGNRPFIAQDDQNAPRV m149.pep ADFYGAEGEIYFKPTPRYRIGVSGDYVRGRLKNLPSLPGREDAYGNRPLIAQADQNAPRV a149 190 200 210 220 230 240 260 270 280 290 PAARLGFHLKASLTDRIDANLDYYRVFAQNKLARYETRTPGHHMLNLGANYRRNTRYGEW m149.pep



The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 601>: g149-1.seq

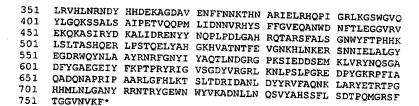
```
1 ATGGCACAAA TCACACTCAA ACCCATTGTT TTATCAATTC TTTTAATCAA
      CACACCCCTC CTCGCCCAAG CGCATGAAAC TGAGCAATCG GTGGGCTTGG
  51
      AAACGGTCAG CGTCGTCGGC AAAAGCCGTC CGCGCGCGAC TTCGGGGCTG
 101
      CTGCACACTT CGACCGCCTC CGACAAAATC ATCTCCGGCG ATACTTTGCG
 151
      CCAAAAAGCC GTCAACTTGG GCGACGCTTT GGACGGCGTA CCGGGCATCC
 201
 251
      ACGCTTCGCA ATACGGCGGC GGCGCATCCG CTCCCGTTAT TCGCGGTCAA
      ACGGGCAGAC GGATTAAAGT ATTGAACCAT CACGGCGAAA CGGGCGATAT
 301
      GGCGGACTTT TCTCCCGATC ACGCCATTAT GGTAGATACC GCCTTGTCGC
 351
 401
      AACAGGTTGA AATCCTGCGC GGGCCGGTTA CGCTCTTGTA CAGCTCGGGC
      AATGTGGCGG GGCTGGTCGA TGTTGCCGAT GGAAAAATCC CCGAAAAAAT
 451
      GCCTGAAAAC GGCGTATCGG GCGaagccgG ATTGCGTTTG AGCAGCGGCA
 501
      ATTTAGAAAA ACTGACATCC GCAGGCATCA ATATCGGACT GGGCAAAAAC
 551
      TTCGTGCTGC ATACCGAAGG CTTGTACCGC AAATCGGGCG ATTACGCCGT
 601
 651
      ACCGCGTTAC CGCAATCTGA AACGCCTGCC CGACAGCCAT GCCGATTCGC
      AAACGGGCAG CATCGGGCTG TCTTGGGTGG GCGAAAAAGG CTTTATCGGC
 701
      GCAGCATACA GCGACCGTCG CGACCGCTAC GGCCTGCCTG CCCACAGCCA
 751
 801
      CGAATACGAT GATTGCCACG CCGACATCAT CTGGCAAAAG AGTTTGATCA
 851
      ACAAACGCTA TTTGCAGCTT TATCCGCACT TGTTGACCGA AGAAGACATC
      GATTACGACA ATCCGGGCTT GAGCTGCGGC TTCCACGACG GCGACGGTGC
 901
      ACACGCACAC ACCCACAACG GCAAACCGTG GATAGACCTG CGCAACAAAC
 951
      GCTACGAACT CCGCGCCGAA TGGAAGCAGC CATTCCCCGG TTTTGAAGCC
1001
      CTGCGCGTAC ATCTGAACCG CAATGACTAC CACCACGACG AAAAAGCAGG
1051
      CGATGCAGTA GAAAACTTCT TCAACAACAA AACACACAAC GCCCGTATCG
1101
1151
      AGTTGCGCCA CCAACCCATA GGCCGTCTGA AAGGCAGCTG GGGCGTGCAA
1201
      TATTTGGGAC AAAAATCCAG CGCGCTTTCC GCCATTCCCG AAACCGTCCA
      ACAACCGATG TTGATTGACA ACAATGTCCG CCATTACAGC TTTTTCGGTG
1251
1301
      TAGAACAGGC AAATTGGGAC AACTTCACGC TTGAAGGCGG CGTACGCGTG
      GAAAAACAAA AAGCCTCCAT CCGGTACGAC AAAGCATTGA TTGATCGAGA
1351
      AAACTACTAC AACCAGCCCC TGCCCGACCT CGGCGCGCAC CGCCAAACCG
1401
      CCCGCTCGTT CGCACTTTCG GGCAACTGGT ATTTCACGCC ACACCACAAA
1451
1501
      CTCAGCCTGA CCGCCTCCCA TCAGGAACGC CTGCCGTCAA CGCAAGAACT
      GTACGCACAC GGCAAGCACG TCGCCACCAA CACCTTTGAA GTCGGCAACA
1551
      AACACCTCAA CAAAGAGCGT TCCAACAATA TCGAACTCGC GCTGGGCTAC
      GAAGGCGACC GCTGGCAATA CAATCTGGCA GCCTACCGCA ACCGATTCGG
1651
      CAACTACATT TACGCCCAAA CCTTAAACGA CGGACGCGGC CCCAAATCCA
1701
      TCGAAGACGA CAGCGAAATG AAGCTCGTGC GCTACAACCA ATCCGGTGCC
1751
      GACTTCTACG GCGCGGAAGG CGAAATCTAC TTCAAACCGA CACCGCGCTA
      CCGCATCGGT GTTTCCGGCG ACTATGTACG AGGCCGTCTG AAAAACCTGC
1851
      CGTCCCTACC CGGCAGGGAA GATCCCTACG GCAAACGTCC CTTCATCGCA
1901
1951
      CAAGCCGACC AAAACGCCCC CCGCATTCCG GCTGCGCGCC TCGGCTTCCA
      CCTGAAAACC TCGCTAACCG ACCGTATCGA TGCCAATTTG GACTACTACC
2001
      GCGTGTTCGC CCAAAACAAA CTCGCCCGCT ACGAAACGCG TACGCCCGGA
2051
2101
      CACCATATGC TCAACCTCGG TGCAAACTAC CGCCGCAATA CGCGCTATGG
      CGAGTGGAAT TGGTACGTCA AAGCCGACAA CCTGCTCAAC CAATCCGTTT
2151
      ACGCCCACAG CAGCTTCCTC TCTGATACGC CGCAAATGGG CCGCAGCTTT
      ACCGGCGGCG TAAACGTGAA GTTTTAA
```

This corresponds to the amino acid sequence <SEQ ID 602; ORF 149-1.ng>: g149-1.pep

1	MAQITLKPIV	LSILLINTPL	LAQAHETEQS	VGLETVSVVG	KSRPRATSGL
51	LHTSTASDKI	ISGDTLRQKA	VNLGDALDGV	PGIHASOYGG	GASAPVIRGO
101	TGRRIKVLNH	HGETGDMADF	SPDHAIMVDT	ALSQQVEILR	GPVTLLYSSG
151	NVAGLVDVAD	GKIPEKMPEN	GVSGEAGLRL	SSGNLEKLTS	AGINIGLGKN
201	<b>FVLHTEGLYR</b>	KSGDYAVPRY	RNLKRLPDSH	ADSOTGSIGL	SWVGEKGFIG
251	AAYSDRRDRY	GLPAHSHEYD	DCHADIIWQK	SLINKRYLOL	YPHLLTEEDI
301	DYDNPGLSCG	<b>FHDGDGAHAH</b>	THNGKPWIDL	RNKRYELRAE	WKOPFPGFEA

BNSDOCID: <WO ___9957280A2 1 >

. ...



# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 603>: m149-1.seq

```
1 ATGGCACAAA CTACACTCAA ACCCATTGTT TTATCAATTC TTTTAATCAA
1 CACACCCCTC CTCGCCCAAG CGCATGAAAC TGAGCAATCG GTGGATTTGG
    51
        AAACGGTCAG CGTCGTCGGC AAAAGCCGTC CGCGCGCCAC GTCGGGGCTG
        TTGCACACTT CGACCGCCTC CGACAAAATC ATCTCCGGCG ATACCTTGCG
        CCAAAAAGCC GTCAACTTGG GCGACGCTTT AGACGGCGTA CCGGGCATCC
  251 ACGCTTCGCA ATACGGCGGC GGCGCGTCTG CTCCCGTCAT TCGCGGTCAA
  301 ACAGGCAGGC GGATTAAAGT GTTGAACCAT CACGGCGAAA CAGGCGATAT
  351 GGCGGATTTT TCGCCCGATC ACGCCATTAT GGTAGATACC GCCTTGTCGC
  401 AACAGGTCGA AATCCTGCGC GGGCCGGTTA CGCTCTTGTA CAGCTCGGGC
       AATGTGGCGG GGCTGGTCGA TGTTGCCGAT GGCAAAATCC CCGAAAAAAT
  451
       GCCTGAAAAC GGCGTATCGG GCGAACTCGG ATTGCGTTTG AGCAGCGGCA
  501
  551
       ATCTGGAAAA ACTCACGTCC GGCGGCATCA ATATCGGTTT GGGCAAAAAC
       TTTGTATTGC ACACGGAAGG GCTGTACCGC AAATCGGGGG ATTACGCCGT
  601
       ACCGCGTTAC CGCAATCTGA AACGCCTGCC CGACAGCCAC GCCGATTCGC
  651
       AAACGGGCAG CATCGGGCTG TCTTGGGTTG GCGAAAAAGG TTTTATCGGC
  701
       GTAGCGTACA GCGACCGTCG CGACCAATAT GGTCTGCCTG CCCACAGCCA
  751
       CGAATACGAT GATTGCCACG CCGACATCAT CTGGCAAAAG AGCTTGATTA
  801
       ACAAACGCTA TTTACAGCTT TATCCGCACC TGTTGACCGA AGAAGACATC
GATTACGACA ATCCGGGCTT GAGCTGCGGC TTCCACGACG ACGATAATGC
  851
  901
       ACACGCACAC ACCCACAGCG GCAGACCGTG GATAGACCTG CGCAACAAAC
  951
       GCTACGAACT CCGTGCCGAA TGGAAGCAAC CGTTCCCCGG TTTTGAAGCC
 1051
       CTGCGCGTAC ACCTGAACCG CAACGACTAC CGCCACGACG AAAAAGCAGG
1101
       CGATGCAGTC GAAAACTTTT TTAACAACCA AACGCAAAAC GCCCGCATCG
       AGTTGCGCCA CCAACCCATA GGTCGTCTGA AAGGCAGCTG GGGCGTGCAA
1151
       TATTTACAAC AAAAATCCAG TGCTTTATCT GCCATATCCG AAGCGGTTAA
1201
       ACAACCGATG CTGCTTGACA ACAAAGTGCA ACATTACAGC TTTTTCGGTG
       TAGAACAGGC AAACTGGGAC AACTTCACGC TTGAAGGAGG CGTACGCGTG
1301
      GAAAAACAAA AAGCCTCCAT TCAGTACGAC AAAGCATTGA TTGATCGGGA
AAACTACTAC AACCACCCCC TGCCCGACCT CGGCGCGCAC CGCCAAACCG
1351
1401
       CCCGCTCATT CGCACTTTCG GGCAACTGGT ATTTCACGCC ACAACACAAA
1451
      CTCAGCCTGA CCGCCTCCCA TCAGGAACGC CTGCCGTCAA CGCAAGAGCT
1501
      GTACGCACAC GGCAAACACG TCGCCACCAA CACCTTTGAA GTCGGCAACA
1551
1601
      AACACCTCAA CAAAGAGCGT TCCAACAATA TCGAACTCGC GCTGGGCTAC
      GAAGGCGACC GCTGGCAATA CAATCTGGCA CTCTACCGCA ACCGCTTCGG
1651
      TAACTACATT TACGCCCAAA CCTTAAACGA CGGACGCGGC CCCAAATCCA
1701
      TCGAAGACGA CAGCGAAATG AAGCTCGTGC GCTACAACCA ATCCGGCGCC
1751
      GACTTCTACG GCGCGGAAGG CGAAATCTAC TTCAAACCGA CACCGCGCTA
1801
      CCGCATCGGC GTTTCCGGCG ACTATGTACG AGGCCGTCTG AAAAACCTGC
1851
      CTTCCCTACC CGGCAGAGAA GATGCCTACG GCAACCGTCC TTTCATCGCA
1901
      CAGGACGACC AAAATGCCCC CCGTGTTCCG GCTGCGCGCC TCGGCTTCCA
1951
      CCTGAAAGCC TCGCTGACCG ACCGTATCGA TGCCAATTTG GACTACTACC
2001
      GCGTGTTCGC CCAAAACAAA CTCGCCCGCT ACGAAACGCG CACGCCCGGA
2051
      CACCATATGC TCAACCTCGG CGCAAACTAC CGCCGCAATA CGCGCTATGG
2101
      CGAGTGGAAT TGGTACGTCA AAGCCGACAA CCTGCTCAAC CAATCCGTTT
2151
      ACGCCCACAG CAGCTTTCTC TCTGATACGC CGCAAATGGG CCGCAGCTTT
2201
2251 ACCGGCGGCG TGAACGTGAA GTTTTAA
```

# This corresponds to the amino acid sequence <SEQ ID 604; ORF 149-1>:

	-				
1	MAQTTLKPIV	LSILLINTPL	LACAHETECS	VDI.ETUCINIC	V00000000
51	LHTSTASDKI	ISGDTLROKA	VNI.GDALDGU	PCIUNCOVCC	KSRPRATSGL
101	TGRRIKVLNH	HGETGDMADF	SDDUATMOOM	PGINASQIGG	GASAPVIRGQ
151	NVAGLVDVAD	GKIPEKMPEN	CACCETCIDI	ALSQQVEILR	GPVTLLYSSG
201	FVLHTEGLYR	KSGDYAVPRY	GARGETGTEF	SSGNLEKLTS	GGINIGLGKN
251	VAYSDRRDOV	CIDAUCUEVA	DOUBBETTER	ADSQTGSIGL	SWVGEKGFIG
301	DADMBCT SCC	GLPAHSHEYD FHDDDNAHAH	DUHADITWQK	SLINKRYLQL	YPHLLTEEDI
351		RHDEKAGDAV	TOTAL OFFICE HEADER	RNKRYELRAE	WKQPFPGFEA
401				ARIELRHOPI	GRLKGSWGVO
451	IDQQKSSALS	AISEAVKQPM	LLDNKVQHYS	FFGVEQANWD	NFTLEGGVRV
	ENGNASIQYD	KALIDRENYY	NHPLPDLCAU	DOWN DOWN TO	CV###
301	PSTASHOEK	LPSTQELYAH	GKHVATNTFE	VGNKHLNKER	SNNIELALGY

	601 651 701 751	DFYGAEGEIY QDDQNAPRVP HHMLNLGANY TGGVNVKF*	FKPTPRYR AARLGFHL	IG VSGD KA SLTD	YVRGRL	KNLPSLPC	RE DAYGN	RPFIA	
m149-1/g149-1 96.2% identity in 758 aa overlap									
m149-	1.pep	MAQTTLK	10 PIVLSILLI	20 NTPLLAQ	3( AHETEQS	SVDLETVSV	40 VGKSRPRA	50 TSGLLHTST	60 ASDKI
g149-	1	111 111		111111	111111	VGLETVSV	11111111	TSGLLHTST	11113
m149-	1.pep	isgDTLR	70 QKAVNLGDAI	80 LDGVPGI	90 HASQYGO	GASAPVIF	.00 KGQTGRRIK	110 VLNHHGETGI	120 DMADE
g149-	1	  ISGDTLR		LDGVPGI 80	  HASQYG0  90	GASAPVIF	 GQTGRRIK .00	 VLNHHGETGI 110	  DMADE   120
m149-	1.pep	SPDHAIM	130 VDTALSQQVI	140 EILRGPV	150 TLLYSS	NVAGLVDV	60 ADGKIPEK	170 MPENGVSGEI	180 LGLRI
g149-	1	SPDHAIM	VDTALSQQVI 130	EILRGPV	TLLYSSO 150	NVAGLVDV	ADGKIPEK 'ADGKIPEK	 MPENGVSGE 170	IIII AGLRI 180
m149-	l.pep	SSGNLEKI	190 LTSGGINIGI	200 LGKNFVL	210 HTEGLYF	KSGDYAVP	20 RYRNLKRL	230 PDSHADSQTO	240 GSIGI
g149-	1	SSGNLEKI	LTSAGINIGI L90	GKNFVL	HTEGLYF 210	KSGDYAVP			  SSIGI  240
m149-	1.pep	SWVGEKG	250 FIGVAYSDRI	260 RDQYGLP	270 AHSHEYD	DCHADIIW	OKSLINKR	290 YLQLYPHLL	300 TEEDI
g149-	1	SWVGEKG	FIGAAYSDRE 250	RDRYGLP	AHSHEYE 270	DCHADIIW	OKSTINKE.		TEEDI 300
m149-	<b>1.pe</b> p	DYDNPGLS			1:11111	RNKRYELR		350 GFEALRVHLN	HILL
g149-	1	DYDNPGLS	SCGFHDGDG# B10	HAHTHNO 320	GKPWIDL 330	RNKRYELR	AEWKQPFP	GFEALRVHLN 350	NRNDY 360
m149-	1.pep	RHDEKAGI	370 DAVENFFNNC	380 TQNARII	390 ELRHQPI	GRLKGSWG	00 VQYLQQKS: 	410 SALSAISEAV	420 KQPM
g149-	1	HHDEKAGI	AVENFFNNE 370	THNARII 380	ELRHQPI 390	GRLKGSWG	VOYLGOKS:	SALSAIPETV 410	/QQPM 420
m149-	1.pep	LLDNKVQH	130 IYSFFGVEQA	440 NWDNFT	450 LEGGVRV	EKOKASIO	60 YDKALIDRI	470 ENYYNHPLPI	480 DLGAH
g149-	1	LIDNNVRH	YSFFGVEQA	NWDNFTI 440	LEGGVRV 450	EKQKASIR	YDKALIDRI 60	illil:     ENYYNQPLPI 470	11111 DLGAH 480
m149-	1.pep	RQTARSFA	190 LSGNWYFTE	500 OHKLSL	510 FASHQER	LPSTQELY	20 AHGKHVATI	530 NTFEVGNKHI	540 INKER
g149-	1	RQTARSFA		:         HHKLSL	IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	LPSTQELY			1111 INKER 540
m149-	1.pep	SNNIELAL	50 GYEGDRWQY	560 NLALYRI	570 NRFGNYI	YAQTLNDG	80 RGPKSIEDI	590 DSEMKLVRYN	600 NOSGA
g149-	1	SNNIELAL	GYEGDRWQY 550	NLAAYRN 560	NRFGNYI 570	YAQTLNDG	 RGPKSIEDI 80		1111 1025GA 000
m149-	l.pep	DFYGAEGE	10 HYFKPTPRY	620 RIGVSGI	630 YVRGRL	KNLPSLPG	40 REDAYGNRI	650 PFIAQDDQNA	660 APRVP
g149-	1	DFYGAEGE	 :IYFKPTPRY :10		 	KNLPSLPG	ii    :i  REDPYGKRI 40	IIIII IIII PFIAQADQNA 650	PRIP 660

551 EGDRWQYNLA LYRNRFGNYI YAQTLNDGRG PKSIEDDSEM KLVRYNQSGA

```
680
                               690
                                      700
          AARLGFHLKASLTDRIDANLDYYRVFAQNKLARYETRTPGHHMLNLGANYRRNTRYGEWN
m149-1.pep
          AARLGFHLKTSLTDRIDANLDYYRVFAONKLARYETRTPGHHMLNLGANYRRNTRYGEWN
g149-1
                       680
                               690
                                      700
                                              710
               730
                       740
                               750
                                     759
m149-1.pep
          WYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKFX
          g149-1
          WYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKFX
                       740
```

## The following partial DNA sequence was identified in N. meningitidis <SEQ ID 605>:

```
1 ATGGCACAAA CTACACTCAA ACCCATTGTT TTATCAATTC TTTTAATCAA
       CACACCCCTC CTCTCCCAAG CGCATGGAAC TGAGCAATCA GTGGGCTTGG
   51
       ANACGGTCAG CGTCGTCGGC ANANGCCGTC CGCGCGCCAC TTCGGGGCTG
  101
       CTGCACACTT CTACCGCCTC CGACAAAATC ATCAGCGGCG ACACCTTGCG
  151
       ACAAAAAGCC GTCAACTTGG GTGATGCTTT AGACGGCGTA CCGGGCATTC
  201
       ATGCCTCGCA ATACGGCGGC GGCGCATCCG CTCCCGTTAT TCGCGGTCAA
  251
       ACAGGCAGAC GGATTAAAGT GTTGAACCAT CACGGCGAAA CGGGCGACAT
  301
       GGCGGACTTC TCTCCAGACC ATGCAATCAT GGTGGACAGC GCCTTGTCGC
  351
       AACAGGTCGA AATCCTGCGC GGTCCGGTTA CGCTCTTGTA CAGCTCGGGC
  401
       AATGTGGCGG GGCTGGTCGA TGTTGCCGAT GGCAAAATCC CCGAAAAAAT
  451
       GCCTGAAAAC GGCGTATCGG GCGAACTCGG ATTGCGTTTG AGCAGCGGCA
  501
       ATCTGGAAAA ACTCACGTCC GGCGGCATCA ATATCGGTTT GGGCAAAAAC
  551
       TTTGTATTGC ACACGGAAGG GCTGTACCGC AAATCGGGGG ATTACGCCGT
  601
       ACCGCGTTAC CGCAATCTGA AACGCCTGCC CGACAGCCAC GCCGATTCGC
  651
       AAACGGGCAG CATCGGGCTG TCTTGGGTTG GCGAAAAAGG CTTTATCGGC
  701
       GCAGCATACA GCGACCGTCG CGACCAATAT GGTCTGCCTG CCCACAGCCA
  751
       CGAATACGAT GATTGCCACG CCGACATCAT CTGGCAAAAG AGTTTGATTA
  801
       ACAAACGCTA TTTGCAGCTT TATCCGCACC TGTTGACCGA AGAAGACATC
  851
      GATTACGACA ATCCGGGCTT GAGCTGCGGC TTTCACGACG ACGATGATGC
  901
      ACACGCCCAT GCCCACAACG GCAAACCTTG GATAGACCTG CGCAACAAAC
  951
1001
      GCTACGAACT CCGCGCCGAA TGGAAGCAAC CGTTCCCCGG TTTTGAAGCC
      CTGCGCGTAC ACCTGAACCG CAACGACTAC CGCCACGACG AAAAAGCAGG
1051
      CGATGCAGTA GAAAACTTTT TTAACAACCA AACGCAAAAC GCCCGTATCG
1101
      AGTTGCGCCA CCAACCCATA GGCCGTCTGA AAGGCAGCTG GGGCGTGCAA
1151
1201
      TATTTGGGAC AAAAATCCAG TGCTTTATCT GCCACATCCG AAGCGGTCAA
      ACAACCGATG CTGCTTGACA ATAAAGTGCA ACATTACAGC TTTTTCGGTG
1251
      TAGAACAGGC AAACTGGGAC AACTTCACGC TTGAAGGCGG CGTACGCGTG
1301
      GAAAAACAAA AAGCCTCCAT CCGCTACGAC AAAGCATTGA TTGATCGGGA
1351
      AAACTACTAC AACCATCCCC TGCCCGACCT CGGCGCGCAC CGCCAAACCG
1401
      CCCGCTCATT CGCACTTTCG GGCAACTGGT ATTTCACGCC ACAACACAAA
1451
      CTCAGCCTGA CCGCCTCCCA TCAGGAACGC CTGCCGTCAA CGCAAGAGCT
1501
      GTACGCACAC GGCAAACACG TCGCCACCAA CACCTTTGAA GTCGGCAACA
1551
1601 AACACCTCAA CAAAGAGCGT TCCAACAATA TCGAACTCGC GCTGGGCTAC
     GAAGGCGACC GCTGGCAATA CAATCTGGCA CTCTACCGCA ACCGCTTCGG
1651
      CAACTACATT TACGCCCAAA CCTTAAACGA CGGACGCGGC CCCAAATCCA
1701
     TCGAAGACGA CAGCGAAATG AAGCTCGTGC GCTACAACCA ATCCGGTGCG
1751
     GACTTCTACG GCGCGGAAGG CGAAATCTAC TTCAAACCGA CACCGCGCTA
1801
     CCGCATCGGC GTTTCCGGCG ACTATGTACG AGGCCGTCTG AAAAACCTGC
1851
     CTTCCCTACC CGGCAGGGAA GACGCCTACG GCAACCGCCC ACTCATTGCC
1901
     CAAGCCGACC AAAACGCCCC TCGCGTTCCG GCTGCGCGCC TCGGCGTCCA
1951
     CCTGAAAGCC TCGCTGACCG ACCGCATCGA TGCCAATTTG GACTACTACC
2001
     GCGTGTTCGC CCAAAACAAA CTCGCCCGCT ACGAAACGCG CACGCCCGGA
2051
     CACCATATGC TCAACCTCGG CGCAAACTAC CGCCGCAATA CGCGCTATGG
2101
     CGAGTGGAAT TGGTACGTCA AAGCCGACAA CCTGCTCAAC CAATCCGTTT
2151
     ACGCCCACAG CAGCTTCCTC TCTGATACGC CGCAAATGGG CCGCAGCTTT
2201
     ACCGGCGGCG TGAACGTGAA GTTTTAA
```

### This corresponds to the amino acid sequence <SEQ ID 606; ORF 149-1.a>:

1	MAQTTLKPIV	LSILLINTPL	LSOAHGTEOS	VCI Emucinio	KSRPRATSGL
	T111 0 1 10 0 D [/ ]	TOGDITIKUKA	VNI.CDAT DCtt	DOTTITE	
101	TGRRIKVLNH	HGETGDMADE	CDUMINATION	PGIHASQYGG	GASAPVIRGQ
151	NVAGLVDVAD FVLHTEGLVB	GKIDEKMDEN	CACCET OF DE	ALSQQVEILR	GPVTLLYSSG
201	FVI.HTEGI.VB	KCCDANADDA	GAZGETGTKT	SSGNLEKLTS	GGINIGLGKN
251	FVLHTEGLYR	CIDAUCUEVA	RNLKRLPDSH	ADSQTGSIGL	SWVGEKGFIG
301	AAYSDRRDQY DYDNPGLSCG	GLPANSHEID	DCHADIIWQK	SLINKRYLQL	YPHLLTEEDI
	DYDNPGLSCG	HAHAUUUUAHAH	AHNGKPWIDL	RNKRYELRAE	WKQPFPGFEA

401 451 501 551 601 701 751 5	LRVHLNRNDY R YLGQKSSALS A' YLGQKSSALS A' YLGQKSSALS A' YLGYKSSALS A' LSLTASHQER L' EGDRWQYNLA L' DFYGAEGEIY FI QADQNAPRVP A' HHMLNLGANY RI TGGVNVKF*	TSEAVKQPM ALIDRENYY PSTQELYAH YRNRFGNYI KPTPRYRIG ARLGVHLKA RNTRYGEWN	LLDNKVQHYS NHPLPDLGAH GKHVATNTFE YAQTLNDGRG VSGDYVRGRL SLTDRIDANL WYVKADNLLN	FFGVEQANWD RQTARSFALS VGNKHLNKER PKSIEDDSEM KNLPSLPGRE DYYRVFAQNK QSVYAHSSFL	NFTLEGGVRV GNWYFTPQHK SNNIELALGY KLVRYNQSGA DAYGNRPLIA LARYETRTPG				
a149-1/m149-1 98.0% identity in 758 aa overlap									
a149-1.pep	10 MAOTTI.KPT		0 30		50 SRPRATSGLLH	60			
m149-1		11111111	11:111 1111	1 111111111	SRPRATSGLLH             SRPRATSGLLH	11111111			
	10	2	0 30		50	FSTASDKI 60			
	70		0 90		110	120			
a149-1.pep				111111111111	GRRIKVLNHHG	1111111			
m149-1	ISGDTLRQKA 70	AVNLGDALDG	VPGIHASQYGG	GASAPVIRGQT	GRRIKVLNHHG 110	ETGDMADF			
		_				120			
a149-1.pep	130 SPDHAIMVDS	SALSQQVEIL	RGPVTLLYSSG	NVAGLVDVADG	170 KIPEKMPENGV	180 SGELGLRL			
m149-1		:		11111111111		11111111			
	130	) 14	0 150		170	180			
	190			220	230	240			
a149-1.pep	SSGNLEKLTS	GGINIGLGK	NFVLHTEGLYR	KSGDYAVPRYR	NLKRLPDSHAD:	SQTGSIGL			
m149-1	SSGNLEKLTS	GGINIGLGK	NFVLHTEGLYR	KSGDYAVPRYR	NLKRLPDSHAD:	SQTGSIGL			
				220	230	240			
a149-1.pep	250 SWVGEKGFIG				290 LINKRYLQLYPI	300			
m149-1		: [ ] [ ] [ ] [ ]	[[[]]]	1111111111		1111111			
11143-1	250	26	270	DCHADIIWQKS: 280	LINKRYLOLYPI 290	HLLTEEDI 300			
	310			340	350	360			
a149-1.pep	DYDNPGLSCG	FHDDDDDAHA	HAHNGKPWIDL	RNKRYELRAEW	KOPFPGFEALR	VHLNRNDY			
m149-1	DYDNPGLSCG	FHDDDNAHAI	HTHSGRPWIDL	RNKRYELRAEWI					
	310	320	330	340	350	360			
a149-1.pep	370 RHDEKAGDAV			400	410 LGQKSSALSATS	420			
m140 1									
m149-1	370	ENFTNNQTQI 380	NARIELRHQPI 390	GRLKGSWGVQYI 400	LQQKSSALSAIS 410	SEAVKQPM 420			
	430	440	450	460	470	480			
a149-1.pep	LLDNKVQHYS	FFGVEQANWI	NFTLEGGVRV	EKOKASIRYDK	ALIDRENYYNHE	T.PDT.GAH			
m149-1	LLDNKVQHYS	FFGVEQANWI	DNFTLEGGVRV	EKQKASIQYDK <i>I</i>	LIDRENYYNHE	PLPDLGAH			
	430	440	450	460	470	480			
a149-1.pep	490			520	530 K <b>HVATNTFEV</b> GN	540			
		1111111111							
m149-1	RQTARSFALS	GNWYFTPQHI	LSLTASHQERI	LPSTQELYAHGI 520	HVATNTFEVGN 530	KHLNKER 540			
	550	560	570	580	590				
a149-1.pep	SNNIELALGY	EGDRWQYNLA	LYRNRFGNYI	AOTLNDGRGP	SIEDDSEMKLV	600 RYNQSGA			
m149-1	SNNIELALGY:		 LYRNRFGNYI		KSIEDDSEMKLV	 RYNOSGA			
	550	560	570	580	590	600			
	610	620	630	640	650	660			





```
a149-1.pep
              \tt DFYGAEGEIYFKPTPRYRIGVSGDYVRGRLKNLPSLPGREDAYGNRPLIAQADQNAPRVP
              m149-1
              DFYGAEGEIYFKPTFRYRIGVSGDYVRGRLKNLPSLPGREDAYGNRPFIAQDDQNAPRVP
                             620
                                               640
                    670
                             680
                                      690
             AARLGVHLKASLTDRIDANLDYYRVFAQNKLARYETRTPGHHMLNLGANYRRNTRYGEWN
  a149-1.pep
              uni manunuminaminaminaminaminamina
             AARLGFHLKASLTDRIDANLDYYRVFAONKLARYETRTPGHHMLNLGANYRRNTRYGEWN
  m149-1
                             680
                                     690
                                              700
                    730
                             740
                                     750
             WYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKFX
 a149-1.pep
             m149-1
             WYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKFX
                    730
                                     750
 The following partial DNA sequence was identified in N.gonorrhoeae <SEQ ID 607>:
 g150.seq
            (partial)
            ..TACTGCAAGG CAGACCCCTT TCCCGCCGCC CTGCTGGCCA ATCAGAAAAT
        1
             CACCGCCCGC CAATCCGATA AAGACGTGCG CCACATCGAA ATCGATTTGA
       51
             GCGGTTCGGA TTTGCACTAC CTCCCGGGCG ACGCGCTCGG CGTTTGGTTT
      101
             GACAACGATC CGGCACTGGT CGGGGAAATC CTAGACCTGC TCGGCATCAA
      151
             TCCGGCAACG GAAATACAGG CGGGCGGAAA AACCCTGCCG GTTGCCTCCG
      201
             CACTGTTATC CCATTTCGAA CTCACGCAAA ACACCCCGC CTTTGTCAAA
      251
             GGCTATGCCA CGTTCGCCGA TAATGACGAA CTCGACCGTA TTGCTGCCGA
      301
             CAACGCCGTT TTGCAAGGCT TTGTGCAAAG CACGCCGATT GCCGGTGTGC
      351
             TGCACCGCTT CCCGGCAAAA CTGACGGCGG AACAATTCGC CGGCCTGCTG
      401
             CGCCCGCTTG CGCCGCGCT GTATTCGATT TCCTCGTCGC AGGCGGAAGC
      451
             GGGGGACGAA GTGCACCTGA CCGTCGGCGC AGTGCGTTTC GAACACGAAG
      501
             GGCGCGCCAG GGCGGCGGC GCATCGGGTT TCTTTGCCGA CCGGCTGGAA
      551
             GAGGACGGCA CGGTGCGCGT GTTTGCGGAA CGCAACGACG GCTTCAGGCT
      601
             GCCCGAAGAC AGCCGCAAGC CGATTGTGAT GATCGGCTCC GGTACCGGCG
      651
             TCGCACCGTT CCGCGCCTTC GTCCAACAAC GTGCCGCAGA AAATGCGGAA
      701
             GGCAGAAACT GGCTGATTTT CGGCAATCCG CATTTTGCCG CCGACTTCCT
      751
             CTATCAGACC GAATGGCAGC AGTTTGCCAA AGACGGCTTC CTGCACAGAT
      801
            ATGACTTCGC CTGGTCGCGC GATCAGGAAG AAAAAATCTA TGTGCAGGAC
     851
            AAAATCCGCG AACAGGCGGA AGGACTTTGG CAATGGCTGC AGGAAGGCGC
     901
            GCATATCTAT GTGTGCGGCG ATGCGGCAAA AATGGCAAAA GAAGTGGAAG
     951
            CCGCCTTGCT GGATGTGATT ATCGGGGCAG GGCATTCGGA CGAAGACGGC
    1001
            GCAGAAGGAT ATTTGGATAT GCTGCGCGAA GAAAAACGCT ATCAGCGTGA
    1051
    1101
            TGTTTATTGA
This corresponds to the amino acid sequence <SEQ ID 608; ORF 150.ng>:
g150.pep
          (partial)
          ..YCKADPFPAA LLANQKITAR QSDKDVRHIE IDLSGSDLHY LPGDALGVWF
       1
            DNDPALVGEI LDLLGINPAT EIQAGGKTLP VASALLSHFE LTQNTPAFVK
      51
            GYATFADNDE LDRIAADNAV LQGFVQSTPI AGVLHRFPAK LTAEQFAGLL
     101
            RPLAPRLYSI SSSQAEAGDE VHLTVGAVRF EHEGRARAGG ASGFFADRLE
     1.51
            EDGTVRVFAE RNDGFRLPED SRKPIVMIGS GTGVAPFRAF VQQRAAENAE
     201
            GRNWLIFGNP HFAADFLYQT EWQQFAKDGF LHRYDFAWSR DQEEKIYVQD
     251
            KIREQAEGLW OWLQEGAHIY VCGDAAKMAK EVEAALLDVI IGAGHSDEDG
     301
            AEGYLDMLRE EKRYQRDVY*
The following partial DNA sequence was identified in N.meningitidis <SEQ ID 609>:
m150.seq
         ATGCAGAACA CAAATCCGCC ATTACCGCCT CTGCCGCCCG AAATCACGCA
      1
         GCTCCTGTCG GGGCTGGACG CGGCACAATG GGCGTGGCTG TCCGGCTACG
      51
         CTTGGGCAAA AGCAGGAAAC GGGGCATCTG CAGGACTGCC CGCGCTTCAG
    151 ACGGCATTGC CGGCGGCAGA ACCTTTTTCC GTAACCGTCC TTTCCGCCTC
    201 GCAAACCGGC AATGCGAAAT CCGTTGCCGA CAAAGCGGCG GACAGCCTGG
         AAGCCGCCGG CATCCAAGTC AGTCGCGCCG AACTGAAAGA CTATAAGGCG
         AAAAACATCG CCGGCGAACG CCGCCTGCTG CTGGTTACCT CCACCCAAGG
    301
         CGAAGGCGAA CCGCCGAAAG AAGCCGTCGT GCTGCACAAA CTGCTGAACG
    351
```

GCAAAAAAGC CCCGAAATTG GACAAACTCC AATTTGCCGT ACTGGGTTTG

401

Automotion of the Commonweal

```
451 GGCGACAGTT CCTATCCGAA TTTCTGTCAG GCAGGTAAAG ATTTCGACCG
     501 GCGTTTTGAA GAATTGGGCG CAAAACGGCT GCTCGAACGC GTTGATGCGG
          ATTTGGACTT TACCGCCTCC GCAAACGCCT GGACAGATAA TATCGCCGCA
     551
          CTCTTAAAAG AAGAAGCCGC AAAAAACCGG GCAACGCCCG CGCCGCAGAC
          AACGCCCCC GCCGGCCTTC AGACGGCACC GGATGGCAGG TACTGCAAGG
          CAGCCCCTT TCCCGCCGCC CTGCTGGCCA ATCAGAAAAT CACCGCCCGC
     701
          CAATCCGATA AAGACGTGCG CCACATCGAA ATCGATTTGA GCGGTTCGGA
     751
     801
          TTTGCACTAC CTCCCGGGCG ACGCGCTCGG CGTTTGGTTT GACAACGATC
          CGGCACTGGT CAGGGAAATC CTAGACCTGC TCGGCATCGA TCCGGCAACG
     851
          GAAATACAGG CGGGCGGAAA GATGATGCCG GTTGCGCGCG CACTTTCATC
     901
          TCATTTCGAA CTCACGCAAA ACACTCCGGC TTTCGTCAAA GGCTATGCCG
     951
          CGTTCGCCCA TTATGAAGAA CTCGATAAAA TCATTGCCGA TAACGCCGTT
    1051 TTGCAGGATT TCGTGCAAAA CACGCCTATT GTCGATGTGC TGCACCGCTT
    1101 CCCGGCAAGC CTGACGGCAG AACAATTCAT CCGTTTACTG CGTCCGCTTG
    1151 CACCCCGTTT GTATTCGATT TCTTCAGCAC AGGCGGAAGT GGGCGATGAA
    1201 GTGCATTTAA CTGTCGGCGT GGTTCGTTTT GAACACGAAG GCCGCGCCAG
    1251 AACGGGCGC GCATCGGGTT TCCTTGCCGA CCGGCTGGAA GAGGACGGCA
    1301 CGGTGCGCGT GTTTGTGGAA CGCAACGACG GCTTCAGGCT GCCCGAAGAC
    1351 AGCCGCAAGC CGATTGTGAT GATCGGCTCG GGCACCGGCG TCGCACCGTT
    1401 CCGCGCTTTC GTCCAACAAC GTGCCGCAGA AAATGCGGAA GGCAAAAACT
1451 GGCTGATTTT CGGCAATCCG CATTTTGCCC GTGATTTTCT CTATCAAACC
         GGCTGATTTT CGGCAATCCG CATTTTGCCC GTGATTTTCT CTATCAAACC
    1501 GAATGGCAGC AGTTTGCCAA AGACGGCTTC CTGCACAGGT ACGATTTCGC
    1551 CTGGTCCCGC GATCAGGAAG AAAAAATCTA TGTGCAGGAC AAAATCCGCG
    1601 AACAGGCGGA AGGACTTTGG CAATGGCTGC AGGAAGGCGC GCATATCTAT
    1651 GTGTGCGGCG ATGCGGCAAA AATGGCAAAA GACGTGGAAG CCGCCTTGCT
    1701 GGATGTGATT ATCGGGGCAG GACATTTGGA CGAAGAGGGC GCAGAAGAAT
    1751 ATTTGGATAT GCTGCGCGAA GAAAAACGCT ATCAGCGTGA TGTTTATTGA
This corresponds to the amino acid sequence <SEQ ID 610; ORF 150>:
m150.pep
          MQNTNPPLPP LPPEITQLLS GLDAAQWAWL SGYAWAKAGN GASAGLPALQ
          TALPAAEPFS VTVLSASQTG NAKSVADKAA DSLEAAGIQV SRAELKDYKA
      51
     101 KNIAGERRLL LVTSTQGEGE PPKEAVVLHK LLNGKKAPKL DKLQFAVLGL
    151 GDSSYPNFCQ AGKDFDRRFE ELGAKRLLER VDADLDFTAS ANAWTDNIAA
    201 LLKEEAAKNR ATPAPQTTPP AGLQTAPDGR YCKAAPFPAA LLANQKITAR
    251 QSDKDVRHIE IDLSGSDLHY LPGDALGVWF DNDPALVREI LDLLGIDPAT
    301 EIQAGGKMMP VARALSSHFE LTQNTPAFVK GYAAFAHYEE LDKIIADNAV
    351 LQDFVQNTPI VDVLHRFPAS LTAEQFIRLL RPLAPRLYSI SSAQAEVGDE
    401 VHLTVGVVRF EHEGRARTGG ASGFLADRLE EDGTVRVFVE RNDGFRLPED
    451 SRKPIVMIGS GTGVAPFRAF VQQRAAENAE GKNWLIFGNP HFARDFLYQT
```

Computer analysis of this amino acid sequence gave the following results:

EWQQFAKDGF LHRYDFAWSR DQEEKIYVQD KIREQAEGLW QWLQEGAHIY VCGDAAKMAK DVEAALLDVI IGAGHLDEEG AEEYLDMLRE EKRYQRDVY*

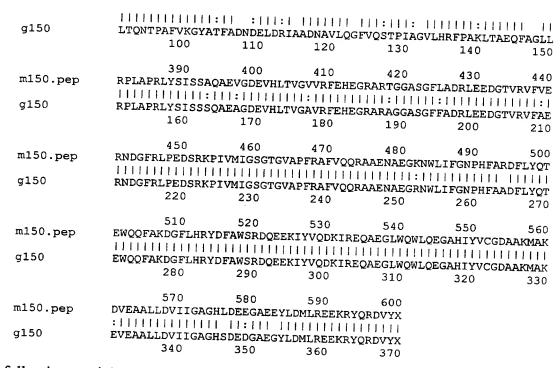
Homology with a predicted ORF from N.gonorrhoeae

ORF 150 shows 91.3% identity over a 369 aa overlap with a predicted ORF (ORF 150.ng) from N. gonorrhoeae:

m150/g150

	210	220	230	240	250	260
m150.pep	LLKEEAAKNRATPA	PQTTPPAGL	QTAPDGRYCK!	APFPAALLAN	QKITARQSDE	KDVRHIE
-150			111			
g150			YCK	ADPFPAALLAN	IQKITARQSDI	KDVRHIE
				10	20	30
	270	280	290	300	310	320
m150.pep	IDLSGSDLHYLPGD	ALGVWFDND				
	- 141111111111			11:111111		
g150	IDLSGSDLHYLPGD	ALGVWFDND	PALVGEILDLI	GINPATEIQA	GGKTLPVASA	ALLSHFE
	40	50	60	70	80	90
	330	340	350	360	370	200
m150.pep						380
mijo.pep	LTQNTPAFVKGYAA	FARYELLDK.	TADNAAFÖDE	VQNTPIVDVL	HRFPASLTAE	COFIRLL

BNSDOCID: <WO___9957280A2_J>



### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 611>:

ı	A TICCA CA A C				
51	COMCOMEN	A CAAATCCGC	ATTACCGCC1	ATGCCGCCC	AAATCACGCA
101	GCTCCTGTC				TCCGGCTACG
151	CTTGGGCAA				CGCGCTTCAG
201	ACGGCATTG			GTAACCGTCC	
	GCAAACCGG			CAAAGCGGCG	
251	AAGCCGCCGC	************		AACTGAAAGA	
301	AAAAACATCO			CTGGTTACCT	
351	CGAAGGCGAZ			GCTGCACAAA	
401	GCAAAAAAG			AATTTGCCGT	
451	GGCGACAGCT	our records		GCGGGCAAAG	
501	ACGTTTTGAZ			GCTCGAACGC	
551	ATTTGGACTT		- 011011000UT	GGACAGATAA	
601	CTCTTAAAAG			GCAACGCCCG	
651	AACGCCCCCC		AGACGGCACC	GGATGGCAGG	
701	CAGACCCCTT		CTGCTGGCCA		CACCGCCCGC
751	CAATCCGATA		CCACATCGAA	ATCGATTTGA	
801	TTTGCACTAC		ACGCGCTCGG	CGTTTGGTTT	GACAACGATC
851	CGGCACTGGT	CAGGGAAATC	CTAGACCTGC	TCGGCATCGA	TCAGGCAACG
901	GAAATACAGG	CGGGCGGAAA	AACCCTGCCG	GTTGCCTCCG	CACTGTTATC
951	CCATTTTGAA	CTCACGCAAA	ACACCCCCCC	CTTTGTCAAA	GGCTATGCCC
1001	CGTTCGCCGA	TGATGACGAA	CTCGACCGTA	TTGCTGCCGA	CAACGCCGTT
1051	TTGCAAGGCT	TTGTGCAAAG	CACGCCGATT	GCCGATGTGC	TGCACCGCTT
1101	CCCGGCAAAA	CTGACAGCGG	AACAATTCGC	CGGCCTACTG	CGCCCGCTTG
1151	CGCCGCGCCT	GTATTCGATT	TCCTCGTCGC	AGGCGGAAGT	GGGGGACGAA
1201	GTGCACCTGA	CCGTCGGCGC	GGTGCGTTTC	GAACACGAAG	GGCGCGCCAG
1251	GGCGGGCGGC	GCATCGGGTT	TCCTTGCCGA	CCGGCTGGAA	GAGGACGGCA
1301	CGGTGCGCGT	GTTTGTGGAA	CGCAACGACG	GCTTCAGGCT	GCCCGAAGAC
1351	AGCCGCAAGC	CGATTGTGAT	GATCGGCTCG	GGCACCGGCG	TCGCACCGTT
1401	CCGCGCTTTC	GTCCAACAAC		AAATGCGGAA	GGCAAAAACT
1451	GGCTGTTTTT	CGGCAATCCG	CATTTTGCCC	GTGATTTTCT	CTATCAAACC
1501	GAATGGCAGC	AGTTTGCCAA			ACGATTTCGC
1551	CTGGTCGCGC	GATCAGGAAG			AAAATCCGCG
1601	AACAGGCGGA			AGGAAGGCGC	
1651	GTGTGCGGCG	ATGCGGCAAA	AATGGCAAAA	GACGTGGAAC	GCATATCTAT
				OFFICO 1 GOVAG	CCGCCTTGCT

1701 1751	GGATGTGATT ATCGGGGCAG GACATTTGGA CGAAGAGGGC GCAGAAGAAAAATTTGGATAT GCTGCGCGAA GAAAAACGCT ATCAGCGTGA TGTTTATTGA	T A						
This corresponds to the amino acid sequence <seq 150.a="" 612;="" id="" orf="">:</seq>								
a150.pep 1 51 101 151 201 251	MQNTNPPLPP MPPEITQLLS GLDAAQWAWL SGYAWAKAGN GASAGLPALO TALPTAEPFS VTVLSASQTG NAKSVADKAA DSLEAAGIQV SRAELKDYKI KNIAGERRLL LVTSTQGEGE PPEEAVVLHK LLNGKKAPKL DKLQFAVLGI GDSSYPNFCR AGKDFDKRFE ELGAKRLLER VDADLDFAAA ADGWTDNIAI LLKEEAAKNR ATPAPQTTPP AGLQTAPDGR YCKADPFPAA LLANQKITAF	A L A R						
301 351 401 451 501	QSDKDVRHIE IDLSGSDLHY LPGDALGVWF DNDPALVREI LDLLGIDQAT EIQAGGKTLP VASALLSHFE LTQNTPAFVK GYAPFADDDE LDRIAADNAV LQGFVQSTPI ADVLHRFPAK LTAEQFAGLL RPLAPRLYSI SSSQAEVGDE VHLTVGAVRF EHEGRARAGG ASGFLADRLE EDGTVRVFVE RNDGFRLPEI SRKPIVMIGS GTGVAPFRAF VQQRAAENAE GKNWLFFGNP HFARDFLYQT EWQQFAKDGF LHRYDFAWSR DQEEKIYVQD KIREQAEGLW QWLQEGAHIY VCGDAAKMAK DVEAALLDVI IGAGHLDEEG AEEYLDMLRE EKRYQRDVY	V E D T						
m150/a150 94.8% iden		60						
m150.pep a150	MQNTNPPLPPLPPEITQLLSGLDAAQWAWLSGYAWAKAGNGASAGLPALQTALPAAEPE	FS						
m150.pep a150	70 80 90 100 110 12 VTVLSASQTGNAKSVADKAADSLEAAGIQVSRAELKDYKAKNIAGERRLLLVTSTQGGG	GE I I						
m150.pep	70 80 90 100 110 12  130 140 150 160 170 18  PPKEAVVLHKLLNGKKAPKLDKLQFAVLGLGDSSYPNFCQAGKDFDRRFEELGAKRLLE   :	30 ER						
a150	PPEEAVVLHKLLNGKKAPKLDKLQFAVLGLGDSSYPNFCRAGKDFDKRFEELGAKRLLE 130 140 150 160 170 18  190 200 210 220 230 24	30						
m150.pep a150	VDADLDFTASANAWTDNIAALLKEEAAKNRATPAPQTTPPAGLQTAPDGRYCKAAPFPA       : :: :	I I						
m150. <b>pe</b> p a150	250 260 270 280 290 30 LLANQKITARQSDKDVRHIEIDLSGSDLHYLPGDALGVWFDNDPALVREILDLLGIDPA	AT     AT						
m150.pep a150	310 320 330 340 350 36 EIQAGGKMMPVARALSSHFELTQNTPAFVKGYAAFAHYEELDKIIADNAVLQDFVQNTP        :	PI PI						
m150.pep a150	370 380 390 400 410 42 VDVLHRFPASLTAEQFIRLLRPLAPRLYSISSAQAEVGDEVHLTVGVVRFEHEGRARTG :	G         						
m150.pep a150	430 440 450 460 470 48 ASGFLADRLEEDGTVRVFVERNDGFRLPEDSRKPIVMIGSGTGVAPFRAFVQQRAAENA !	E E						
m150.pep	490 500 510 520 530 54 GKNWLIFGNPHFARDFLYQTEWQQFAKDGFLHRYDFAWSRDQEEKIYVQDKIREQAEGL	w						



GKNWLFFGNPHFARDFLYQTEWQQFAKDGFLHRYDFAWSRDQEEKIYVQDKIREQAEGLW a150 490 500 510 520 560 570 580 590 QWLQEGAHIYVCGDAAKMAKDVEAALLDVIIGAGHLDEEGAEEYLDMLREEKRYQRDVYX m150.pep a150 QWLQEGAHIYVCGDAAKMAKDVEAALLDVIIGAGHLDEEGAEEYLDMLREEKRYQRDVYX 550 560 570 580 590

```
The following partial DNA sequence was identified in N.gonorrhoeae <SEQ ID 613>:
```

```
ATGAAACAAA TCCGCAACAT CGCCATCATC GCACACGTCG ACCACGGCAA
    7
       AACCACATTG GTCGACCAAC TGCTGCGCCA ATCCGGCACA TTCCGCGCCA
   51
       ACCAGCAGGT TGACGAGCGC GTGATGGACA GCAACGACCT TGAAAAAGAA
  101
       CGCGGCATCA CCATCCTCGC CAAAAACACC GCCATCGATT ACGAAGGCTG
  151
       CCACATCAAT ATCGTCGACA CGCCGGGACA CGCCGACTTC GGCGGCGAAG
  201
       TGGAGCGCGT TTTGGGGATG GTGGATTGCG TCGTCTTGTT GGTGGACGCA
  251
      CAGGAAGGTC CGATGCCGCA AACCCGTTTC GTGACCAAAA AAGCCTTGGC
  301
  351
       TTTGGGGCTG AAACCGATTG TCGTCATCAA CAAAATCGAC AAACCGTCCG
      CCCGTCCGAG CTGGGTTATC GACCAGACTT TCGAGTTGTT CGACAACTTG
      GGTGCGACCG ACGAGCAGTT GGATTTCCCG ATTGTTTACG CTTCAGGTTT
      GAGCGGCTTT GCCAAGCTGG AAGAAACCGA CGAGAGCAGC GATATGCGCC
  501
      CGCtgttcgA CACCATCCTA AAATACAcgc ctgCACCGAG CGGCAGCGCG
 551
      GACGAGCCGC TGCAACTGCA AATTTCCCAA CTCGACTACG ACAACTACAC
  601
      CGGCCGCCTC GGTATCGGTC GTATCTTGAA CGGACGCATC AAACCCGGCC
  651
      AAACCGTTGC CGTGATGAAC CACGAGCAGC AAATCGCCCA AGGCCGCATC
  701
      AACCAGCTTT TGGGTTTCAA AGGCTTGGAA CGCGTGCCGC TTGAAGAAGC
  751
      CGAAGCCGGC GACATTGTGA TTATTTCCGG TATCGAAGAC ATCGGCATCG
 801
 851
      GCGTAACCAT CACCGACAAA GACAACCCCA AAGGCCTGCC GATGTTGAGC
      GTGGACGAAC CGACGCTGAC GATGGACTTT ATGGTAAACA CCAGCCCGCT
 901
      CGCAGGTACA GAAGGCAAAT TCGTGACCAG CCGCCAAATC CGCGACCGCC
      TGCAAAAAGA ATTGCTGACC AACGTTGCCC TGCGCGTGGA AGACACCGCC
1001
      GatgCCGACG TGTTCCGCGT ATCCGGGCGC GGCGAACTGC ACCTGACGAT
1051
      TTTGCTGGAA AATATGCGCC GCGAAGGCTA CGAACTCGCC GTCGGCAAGC
1101
      CGCGCGTCGT GTACCGAGAC ATCGACGGTC AAAAATGCGA ACCTTATGAA
1151
      AACCTGACTG TGGACGTACC CGaCGaCAAC CAAGGCGCGG TAATGGAAGA
      ACTCGGCCGC CGCCGTGGCG AACTGACCAA TATGGAAAGC GACGGCAACG
1251
      GacgCACCCG CCTCGAATAC CATATTCCAG CGCGCGGCTT GATCGGTTTC
1301
      CAAGGCGAAT TCATGACCCT GACGCGCGGC GTCGGGCTGA TGAGCCaCGT
1351
1401 GTTcgacgac tacgegcccg tcaAACCCGA TATGCCCGGC CGCCACAACG
     GCGTactggt GtcccaAGAG CAGGGCGAGG CGGTTGCTTA CGCCTTGTGG
1501 AATCTTGAAG ACCGCGGCCG TATGTTCGTA TCGCCCAACG ACAAAATCTA
1551 CGAAGGTATG ATTATCGGCA TCCACAGCCG CGACAACGAT TTGGTGGTCA
1601 ACCCGCTCAA AGGCAAAAAA CTCACCAATA TCCGTGCCAG CGGTACCGAC
1651 GAAGCGGTGC GCCTGACCAC GCCGATCAAA CTGACGCTGG AAGGCGCGGT
     CGAGTTTATC GACGATGACG AGCTGGTGGA AATCACGCCG CAAtccatcc
     gcctgcgcat gcgttacctG AGCGaattgg aacgccgccg tcaTTTTAAA
1801
      AagctgGATT AA
```

# This corresponds to the amino acid sequence <SEQ ID 614; ORF 151.ng>: g151.pep

1	MKQIRNIAII	AHVDHGKTTL	VDQLLRQSGT	FRANOOVDER	VMDSNDLEKE
51	RGITILAKNT	AIDYEGCHIN	IVDTPGHADF	GGEVERVI.CM	MOMALIANA
101	QEGPMPQTRF	VTKKALALGL	KPIVVINKID	KPSARPSWVT	DOTERT POM
151	GATDEQLDFP	IVYASGLSGF	AKLEETDESS	DMRPLEDTII.	KVTDADCCCA
201	DELLÖLÖISÖ	LDYDNYTGRL	GIGRILNGRI	KPGOTVAVMN	HECOTAGERT
251	NQLLGFKGLE	RVPLEEAEAG	DIVIISGIED	TGTGVTTTDK	DNDVCI DNI C
301	VDEPTLIMDF	MVNTSPLAGT	EGKFVTSROI	RDRLOKELLT	MVAT.DUEDEN
351	DADVFRVSGR	GELHLTILLE	NMRREGYELA	VCKDDIMVDD	TDCOVCEDVE
401	NTIADABDDN	QGAVMEELGR	RRGELTNMES	DGNGRTRI.EV	WIDARCI TOR
451	QGEFMTLTRG	VGLMSHVFDD	YAPVKPDMPG	RHNGVLVSQE	QGEAVAYALW

그 121 1세. . 그 시 그 등 일

```
501 NLEDRGRMFV SPNDKIYEGM IIGIHSRDND LVVNPLKGKK LTNIRASGTD
         EAVRLTTPIK LTLEGAVEFI DDDELVEITP QSIRLRMRYL SELERRRHFK
     551
     601 KLD*
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 615>:
m151.seq
         ATGAAACAAA TCCGCAACAT CGCCATCATC GCCCACGTCG ACCACGGCAA
      1
         AACCACATTG GTCGACCAAC TGCTGCGCCA ATCCGGCACA TTCCGCGCCA
      51
         ACCAGCAGGT TGACGAGCGC GTGATGGACA GCAACGACCT TGAAAAAGAA
     151 CGCGGCATCA CCATCCTCGC CAAAAACACC GCCATCGATT ACGAAGGCTA
     201 CCACATCAAT ATCGTCGACA CGCCGGGACA CGCCGACTTC GGCGGCGAAG
     251 TAGAGCGCGT TTTGGGGATG GTGGACTGCG TCGTCTTGTT GGTGGACGCG
     301 CAGGAAGGCC CGATGCCGCA AACCCGTTTC GTGACCAAAA AAGCCTTGGC
    351 TTTGGGGCTG AAACCGATTG TCGTCATCAA CAAAATCGAC AAGCCGTCCG
     401 CTCGTCCGAG CTGGGTTATC GACCAAACTT TCGAGCTGTT CGACAACTTG
     451 GGCGCGACCG ACGAGCAGTT GGATTTCCCG ATTGTTTACG CTTCAGGGTT
     501 GAGCGGTTTC GCCAAATTGG AAGAAACCGA CGAGAGCAAC GACATGCGTC
     551 CGCTGTTCGA TACTATCTTA AAATATACGC CTGCACCGAG CGGCAGCGCG
     601 GACGAAACGC TGCAACTGCA AATTTCCCAA CTCGACTACG ACAACTACAC
     651 CGGCCGCCTC GGTATCGGTC GTATCTTGAA CGGACGCATC AAACCCGGCC
     701 AAACCGTTGC CGTCATGAAC CACGATCAGC AAATCGCCCA AGGCCGCATC
     751 AACCAGCTTT TGGGTTTCAA AGGTTTGGAA CGCGTGCCGC TTGAAGAAGC
         CGAAGCCGGC GACATCGTGA TTATTTCCGG TATCGAAGAC ATCGGTATCG
         GCGTAACCAT CACCGACAAA GACAATCCCA AAGGCCTACC GATGTTGAGC
          GTGGACGAAC CGACGCTGAC GATGGACTTT ATGGTCAACA CCAGCCCGCT
         GGCGGGTACG GAAGGCAAAT TCGTAACCAG CCGCCAAATC CGCGACCGCC
         TGCAAAAAGA ATTGCTGACC AACGTCGCCC TGCGCGTGGA AGATACCGCC
    1001
    1051 GATGCCGACG TGTTCCGCGT ATCCGGGCGC GGCGAGCTGC ACCTGACCAT
    1101 TTTGCTGGAA AACATGCGCC GCGAAGGCTA CGAACTCGCC GTCGGCAAAC
    1151 CGCGCGTCGT GTACCGCGAC ATCGACGGTC AAAAATGCGA ACCGTATGAA
    1201 AACCTGACCG TGGATGTACC CGACGACAAC CAAGGCGCGG TAATGGAAGA
    1251 ACTCGGCCGC CGCCGTGGCG AACTGACTAA TATGGAAAGC GACGGCAACG
    1301 GACGCACCCG CCTCGAATAC CATATTCCAG CGCGCGGCTT GATCGGTTTC
    1351 CAAGGCGAAT TTATGACCCT GACGCGGGG GTCGGGCTGA TGAGCCACGT
    1401 GTTCGACGAT TACGCGCCCG TCAAACCCGA TATGCCCGGC CGCCACAACG
    1451 GCGTGCTGGT GTCCCAAGAG CAGGGCGAGG CAGTCGCTTA CGCCTTGTGG
    1501 AATCTGGAAG ACCGCGGCCG TATGTTCGTA TCGCCCAACG ACAAAATCTA
         CGAAGGCATG ATTATCGGCA TCCACAGTCG CGACAACGAT TTGGTGGTCA
    1551
    1601 ACCCGCTCAA AGGCAAAAA CTTACCAACA TCCGTGCCAG CGGTACCGAC
    1651 GAAGCCGTTC GCCTGACCAC GCCAATCAAG CTGACGCTGG AAGGTGCGGT
    1701 TGAGTTTATC GACGATGACG AACTCGTTGA AATCACGCCG CAATCCATCC
    1751 GTCTGCGCAA GCGTTACTTG AGCGAATTGG AACGCCGCCG CCACTTTAAA
    1801 AAGCTGGATT GA
This corresponds to the amino acid sequence <SEQ ID 616; ORF 151>:
m151.pep
         MKQIRNIAII AHVDHGKTTL VDQLLRQSGT FRANQQVDER VMDSNDLEKE
      51 RGITILAKNT AIDYEGYHIN IVDTPGHADF GGEVERVLGM VDCVVLLVDA
     101 QEGPMPQTRF VTKKALALGL KPIVVINKID KPSARPSWVI DQTFELFDNL
     151 GATDEQLDFP IVYASGLSGF AKLEETDESN DMRPLFDTIL KYTPAPSGSA
     201 DETLQLQISQ LDYDNYTGRL GIGRILNGRI KPGQTVAVMN HDQQIAQGRI
     251 NQLLGFKGLE RVPLEEAEAG DIVIISGIED IGIGVTITDK DNPKGLPMLS
     301 VDEPTLTMDF MVNTSPLAGT EGKFVTSRQI RDRLQKELLT NVALRVEDTA
     351 DADVFRVSGR GELHLTILLE NMRREGYELA VGKPRVVYRD IDGQKCEPYE
     401 NLTVDVPDDN QGAVMEELGR RRGELTNMES DGNGRTRLEY HIPARGLIGF
     451 QGEFMTLTRG VGLMSHVFDD YAPVKPDMPG RHNGVLVSQE QGEAVAYALW
     501 NLEDRGRMFV SPNDKIYEGM IIGIHSRDND LVVNPLKGKK LTNIRASGTD
          EAVRLTTPIK LTLEGAVEFI DDDELVEITP QSIRLRKRYL SELERRRHFK
```

Computer analysis of this amino acid sequence gave the following results:

BNSDOCID: <WO___9957280A2_I_>

Homology with a predicted ORF from N.gonorrhoeae ORF 151 shows 99.2% identity over a 603 aa overlap with a predicted ORF (ORF 151.ng) from N. gonorrhoeae:

m151/g151

m151.pep	10	20	30	40	50	60
zor.pcp	MKQIRNIAIIAHV	DHCKILTADO	LLRQSGTFRI	ANQQVDERVMD.	SNDLEKERGI	ITILAKNT
g151		DHGKTTLVDO	LLROSGTER			
	10	20	30	40	50 50	60 TILLAKNI
	7.0			-		00
m151.pep	70	80 TPCHADECCE!	90	100	110	120
~ -	AIDYEGYHINIVD					
g151	AIDYEGCHINIVD	TPGHADFGGE	VERVLGMVDC	VVLLVDAOEGE		ון ן ן ן ן ן
	70	80	90	100	110	ALALGE 120
	130	1.40				120
m151.pep	KPIVVINKIDKPSA	140 ARPSWYIDOT:	150	160	170	180
g151		ARPSWVIDQTE	ELFDNLGAT	DEOLDFPIVYA	SGLSGFAKI.	:        2247THT
	130	140	150	160	170	180
	190	200	210			
m151.pep	DMRPLFDTILKYTE	PAPSGSADETT	OLOTSOLDY	220 DNYTCRICTOR	230	240
- 1 5 1						
g151	DMRPLFDTILKYTF	HIDGOMDEFI	CrOT2OTDA	DNYTGRLGIGR	ILNGRIKPG	NMVAVTC
	190	200	210	220	230	240
	250	260	270	280	290	200
m151.pep	HDQQIAQGRINQLL	GFKGLERVPL	EEAEAGDTU	TECTEDICIO		300 CGT.PMT.S
g151						
9101	HEQQIAQGRINQLL 250	GFKGLERVPL 260	EFAFAGDIV.	LISGIEDIGIG	VTITDKDNP	KGLPMLS
		200	270	280	290	300
m151.pep	310	320	330	340	350	360
misi.pep	VDEPTLTMDFMVNT.	SPLAGTEGKE	VTSRQIRDRI	QKELLTNVAL		
g151	VDEPTLTMDFMVNT					
	310	320	330	340	(VEDTADADV 350	
	27.0			340	330	360
m151.pep	370 GELHLTTLLENMODE	380	390	400	410	420
	GELHLTILLENMRRI	:	RVVYRDIDGO	KCEPYENLTVD	VPDDNQGAV	MEELGR
g151	GELHLTILLENMRRE	EGYELAVGKPF	RVVYRDIDGO	TITLETTELLER KCEPYENLTVD		MEELCD
	370	380	390	400	410	MEELGR 420
	430	440				
m151.pep	RRGELTNMESDGNGF	44U STRLEYHIDAR	450	460	470	480
g151		/IVDDIUTEWE	GLIGFQGEF	MTLTRGVGLMS	HVFDDYAPV:	KPDMPG
	430	440	450	460	470	480
	490	500	510	520	500	
m151.pep	RHNGVLVSQEQGEAV	AYALWNLEDR	CRMEVEDNIN	VIVECNIT CITY	530 SPONDI 177011	540
g151						
9131	RHNGVLVSQEQGEAV	TITEDE	GRMF AS SUDI	KIYEGMIIGIH:	SRDNDLVVNI	LKGKK
	490	500	510	520	530	540
	550	560	570	580	590	600
m151.pep	LTNIRASGTDEAVRL	TTPIKLTLEG	AVEETODDET	VETEROGERE	37FD117 0	600 RRRHFK
g151						
=	LTNIRASGTDEAVRL	560	AVEFIDDDEI 570	VEITPQSIRLE	RMRYLSELER	RRHFK
			370	580	590	600

m151.pep KLDX  $\Pi\Pi\Pi$ g151 KLDX The following partial DNA sequence was identified in N. meningitidis <SEQ ID 617>: a151.seq ATGAAACAAA TCCGCAACAT CGCCATCATC GCCCACGTCG ACCACGGCAA 1 51 AACCACATTG GTCGACCAAC TGCTGCGCCA ATCCGGCACA TTCCGCGCCA 101 ACCAGCAGGT TGACGAGCGC GTGATGGACA GCAACGACCT TGAAAAAGAA CGCGGCATCA CCATCCTCGC CAAAAACACC GCCATCGATT ACGAAGGCTA 201 CCACATCAAT ATCGTCGACA CGCCGGGACA CGCCGACTTC GGCGGCGAAG 251 TAGAGCGAGT TTTGGGGATG GTGGACTGCG TCGTCTTGTT GGTGGACGCG 301 CAGGAAGGCC CGATGCCGCA AACCCGTTTC GTGACCAAAA AAGCCTTGGC 351 TTTGGGGCTG AAACCGATTG TCGTCATCAA TAAAATCGAC AAACCGTCCG 401 CCCGTCCGAG CTGGGTCATC GACCAAACTT TCGAGCTGTT CGACAACTTG 451 GGCGCGACCG ACGAGCAGTT GGATTTCCCG ATTGTTTATG CTTCCGGTCT 501 GTCCGGTTTC GCCAAATTGG AAGAAACCGA CGAGAGCAAC GACATGCGTC 551 CGCTGTTCGA TACTATCTTA AAATATACGC CTGCACCGAG CGGCAGCGCG 601 GACGAAACGC TGCAACTGCA AATTTCCCAA CTCGACTACG ACAACTACAC 651 CGGCCGCCTC GGTATCGGTC GTATCTTGAA CGGACGTATC AAGCCCGGTC 701 AAGTTGTTGC CGTCATGAAC CACGATCAAC AAATCGCCCA AGGCCGCATC 751 AACCAGCTTT TGGGTTTCAA AGGTTTAGAA CGCGTGCCGC TTGAAGAAGC 801 CGAAGCCGGC GACATCGTGA TTATTTCCGG TATTGAAGAC ATCGGCATCG 851 GCGTAACCAT CACCGACAAA GACAACCCCA AAGGCCTGCC GATGTTGAGC GTGGACGAAC CGACGCTGAC GATGGACTTT ATGGTCAACA CCAGCCCGTT 951 GGCAGGTACG GAAGGCAAAT TCGTAACCAG CCGCCAAATC CGCGACCGCC 1001 TGCAAAAAGA ATTGCTGACC AACGTCGCCC TGCGCGTGGA AGATACCGCC 1051 GATGCCGACG TGTTCCGCGT ATCCGGGCGC GGCGAGCTGC ACCTGACCAT 1101 TTTGCTGGAA AACATGCGCC GCGAAGGCTA CGAACTCGCC GTCGGCAAAC CGCGCGTCGT GTACCGCGAC ATCGACGGTC AAAAATGCGA ACCGTATGAA 1151 1201 AACCTGACCG TGGACGTACC CGACGACAAC CAAGGCGCGG TAATGGAAGA 1251 ACTCGGCCGC CGCCGTGGCG AACTGACTAA TATGGAAAGC GACGGCAACG 1301 GACGCACCCG CCTCGAATAC CATATTCCAG CGCGCGGCTT GATCGGCTTC 1351 CAAGGCGAAT TTATGACCCT GACGCGCGGG GTCGGGCTGA TGAGCCACGT 1401 GTTCGACGAT TACGCGCCCG TCAAACCCGA TATGCCTGGC CGCCACAACG 1451 GCGTGCTGGT GTCCCAAGAG CAGGGCGAGG CAGTCGCTTA CGCCTTGTGG 1501 AATCTGGAAG ACCGCGGCCG TATGTTCGTA TCGCCCAACG ACAAAATCTA CGAAGGTATG ATTATCGGCA TCCACAGTCG CGACAACGAT TTGGTGGTCA 1551 ACCCGCTCAA AGGCAAAAAA CTTACCAACA TCCGTGCCAG CGGTACCGAC GAAGCCGTTC GCCTGACCAC GCCGATTAAG CTGACGCTGG AAGGTGCGGT 1651 1701 CGAGTTTATC GACGATGATG AGCTGGTAGA AATCACGCCG CAATCCATCC GTCTGCGCAA GCGTTACTTG AGCGAATTGG AACGCCGCCG CCATTTCAAA AAGCTAGATT GA This corresponds to the amino acid sequence <SEQ ID 618; ORF 151.a>: a151.pep MKQIRNIAII AHVDHGKTTL VDQLLRQSGT FRANQQVDER VMDSNDLEKE RGITILAKNT AIDYEGYHIN IVDTPGHADF GGEVERVLGM VDCVVLLVDA QEGPMPQTRF VTKKALALGL KPIVVINKID KPSARPSWVI DQTFELFDNL 101 GATDEQLDFP IVYASGLSGF AKLEETDESN DMRPLFDTIL KYTPAPSGSA DETLQLQISQ LDYDNYTGRL GIGRILNGRI KPGQVVAVMN HDQQIAQGRI 251 NQLLGFKGLE RVPLEEAEAG DIVIISGIED IGIGVTITDK DNPKGLPMLS VDEPTLTMDF MVNTSPLAGT EGKFVTSRQI RDRLQKELLT NVALRVEDTA 351 DADVFRVSGR GELHLTILLE NMRREGYELA VGKPRVVYRD IDGQKCEPYE 401 NLTVDVPDDN QGAVMEELGR RRGELTNMES DGNGRTRLEY HIPARGLIGF QGEFMTLTRG VGLMSHVFDD YAPVKPDMPG RHNGVLVSQE QGEAVAYALW NLEDRGRMFV SPNDKIYEGM IIGIHSRDND LVVNPLKGKK LTNIRASGTD 551 EAVRLTTPIK LTLEGAVEFI DDDELVEITP QSIRLRKRYL SELERRRHFK 601

20

30

40

50

60

BNSDOCID: <WO_ 9957280A2 1 >

KLD*

m151/a151 99.8% identity in 603 aa overlap

Light Color of the All Aska

m151.pep	MKQIRNIAIIAHV	DHGKTTLVDC	T.T.DOCCTRD	. NOOTTO TO THE		
a151						
g121	MKQIRNIAIIAHVI	DHGKTTLVDQ 20	LLROSGTFRA 30	MQQVDERVMD	SNDLEKERGI	TILAKNT
		20	30	40	50	60
m151.pep	70 AIDYEGYHTNTVD	80	90	100	110	120
	AIDYEGYHINIVD					
a151	AIDYEGYHINIVDT	PGHADFGGE 80	A TRATEMADO	VVLLVDAQEG	PMPQTRFVTK	KALALGL
	, 0	80	90	100	110	120
m151.pep	130 KPIVVINKIDEDED	140	150	160	170	180
	KPIVVINKIDKPSA					
a151	KPIVVINKIDKPSA 130	WESMATDOL!	FELFUNLGAT	DEOLDFPIVY	ASGLSGFAKL	EETDESN
	130	140	<b>1</b> 50	160	170	180
m151.pep	190	200	210	220	230	240
mror.pep	DMRPLFDTILKYTP					
a151		WE SOSMDETT	CLQISQLDY	DNYTGRLGIGE		
	190	200	210	220	230	240
m151 man	250	260	270	280	290	300
ml51.pep	HDQQIAQGRINQLL	GFKGLERVPI	EEAEAGDIV	ISGIEDIGIG	VTITDKDNPK	
a151		TI MODERA & T	EEAEAGDIV	ISGIEDIGIG	 VTITDKDNPK	GLDMLS
	250	260	270	280	290	300
-151	310	320	330	340	350	360
m151.pep	VDEPTLTMDFMVNTS	PLAGTEGKE	VTSRQIRDRI	OKELLTNVAL	DUEDES	
a151		PLAGTEGKE	VTSRQIRDRL			 EDUSCE
	310	320	330	340	350	360
-151	370	380	390	400	410	420
m151.pep	GELHLTILLENMRRE	GYELAVGKP	RVVYRDIDGQ	KCEPYENLTVI		
a151	GELHLTILLENMRRE	GIELAVGKE	RVVYRDIDGQ		  VPDDNOGAVI	HIIIII MEELCR
	370	380	390	400	410	420
-151	430	440	450	460	470	480
m151.pep	RRGELTNMESDGNGR	TRLEYHIPAI	RGLIGFOGEF	MTLTRGVGLMS		
a151		TUTEIUT LAF	RGLIGFQGEF	MTLTRGVGLMS		(PDMPC
	430	440	450	460	470	480
-151	490	500	510	520	530	540
m151.pep	RHNGVLVSQEQGEAV	AYALWNLEDR	RGRMFVSPNDI	(IYEGMIIGIH		
a151		AYALWNLEDR	RGRMFVSPNDE			ILCAR
	490	500	510	520	530	540
m151	550	560	570	580	590	600
m151.pep	LTNIRASGTDEAVRLT	TPIKLTLEG	AVEFIDDDEL	VETEROOFT		
a151	LTNIRASGTDEAVRLT					
	550	560	570	580	590	600
1.5.4						
m151.pep	KLDX					
a151	KLDX					

The following partial DNA sequence was identified in N.gonorrhoeae <SEQ ID 619>: g152.seq

- 1 ATGAAAAaca aAACCaaagt ctgGGacttc cCcacccgcc ttTTCCactG
  - 51 GctgcttgCC gCATCCctgc CCTTTATGTG gtatagCGCA AAAGCCGGCG
- 101 GCGataTGCT GCaatgGCAC ACGCGCGTCG GGCTGCTCGT CCTTTTCCTG

```
151 CTCGTATTCC GCCTCTGCTG GGGCATTTGG GGCAGCGATA CCGCCCGTTT
     201 CTCccgTtTC GTCCGAGGTT GGGCAGGTAT ACGCGGCTAT CTGAAAAAcg
     251 gCATTCCCGA ACAtatcCAG CCCGGACACA ACCCCTTGGG CGCACTgatg
     301
          gtcGTTGCGC TTTTGgccgc cgtcTCATTT CAagtcggcA CGGGGCTTTT
          Tgccgccaat gaaaacacct tcagcaCCAa cggctacctc aaccatttgg
     351
          tttccgaaca tacgGGCAGC CTTATACGGA AAATCCACCT CAACTTTTTC
          AAGCTGCTCG CCGTTTTTTC CGCAGTCCAC ATCGCCGCCG TCGCCGCATA
     451
     501 CCGCATATTC AAAAAGAAAA ACCTCGTCCG CCCGATGATA ACCGGCTTCA
     551 AATACATCGA AGGCAAAACC TCAATCCGCT TTGCCGGCAA AGCCGCGCTT
         GCCGCCGCAT TATCGGTTGC CGCGCTTGCC GCAGCCGCCA TCCTGCTCCT
     601
          GTCCTGA
This corresponds to the amino acid sequence <SEO ID 620; ORF 152.ng>:
q152.pep
          MKNKTKVWDF PTRLFHWLLA ASLPFMWYSA KAGGDMLQWH TRVGLLVLFL
      51
          LVFRLCWGIW GSDTARFSRF VRGWAGIRGY LKNGIPEHIQ PGHNPLGALM
     101
          VVALLAAVSF QVGTGLFAAN ENTFSTNGYL NHLVSEHTGS LIRKIHLNFF
          KLLAVFSAVH IAAVAAYRIF KKKNLVRPMI TGFKYIEGKT SIRFAGKAAL
     151
     201
          AAALSVAALA AAAILLLS*
The following partial DNA sequence was identified in N. meningitidis <SEQ ID 621>:
m152.seq
      1
         ATGAAAAACA AAACCAAAGT CTGGGACCTC CCCACCCGCC TTTTCCACTG
         GCTGCTTGCC GCGTCCCTGC CCTTTATGTG GTATAGCGCG AAAGCCGGCG
      51
         GCGATATGCT GCAATGGCAC ACGCGCGTCG GGCTGTTCGT CCTTTTCCTG
         CTCGTATTTC GCCTCTGCTG GGGCATTTGG GGCAGCGATA CCGCCCGTTT
     201 TTCCCGTTTC GTCCAAGGCT GGGCAGGCAT ACGCGGCTAT CTGAAAAACG
     251 GTATTCCCGA ACACATCCAG CCCGGACACA ACCCCTTGGG CGCACTGATG
     301 GTCGTTGCGC TTTTGGCCGC CGTGTCCTTC CAAGTCGGCA CCGGGCTTTT
     351 TGCCGCCGAT GAAAACACCT TCAGCACCAA CGGCTACCTC AACCATTTGG
     401 TTTCCGAACA TACGGGCAGC CTTATGCGGA AAATCCACCT CAACTTTTTC
     451 AAGCTGCTCG CCGTTTTTC TGCAATCCAC ATCGCCGCCG TCGCCGCATA
     501 CCGCGTATTC AAAAAGAAAA ACCTCATCCT CCCGATGATA ACCGGCTTCA
     551 AATACATCGA AGGCAAAACC TCAATCCGCT TTGCAGGCAA AGCCGCGCTT
     601 GCCGCCGCAT TATCGGTTGC CTCGCTTGCC GCAGCCGCCA TCCTGCTCCT
     651 GTCCTGA
This corresponds to the amino acid sequence <SEQ ID 622; ORF 152>:
m152.pep
         MKNKTKVWDL PTRLFHWLLA ASLPFMWYSA KAGGDMLOWH TRVGLFVLFL
      1
      51
         LVFRLCWGIW GSDTARFSRF VQGWAGIRGY LKNGIPEHIQ PGHNPLGALM
     101
         VVALLAAVSF OVGTGLFAAD ENTFSTNGYL NHLVSEHTGS LMRKIHLNFF
     151
         KLLAVFSAIH IAAVAAYRVF KKKNLILPMI TGFKYIEGKT SIRFAGKAAL
         AAALSVASLA AAAILLLS*
Computer analysis of this amino acid sequence gave the following results:
Homology with a predicted ORF from N.gonorrhoeae
ORF 152 shows 95.4% identity over a 218 aa overlap with a predicted ORF (ORF 152.ng)
from N. gonorrhoeae:
m152/g152
                    10
                              20
                                        30
                                                  40
                                                           50
m152.pep
            {\tt MKNKTKVWDLPTRLFHWLLAASLPFMWYSAKAGGDMLQWHTRVGLFVLFLLVFRLCWGIW}
            g152
            MKNKTKVWDFPTRLFHWLLAASLPFMWYSAKAGGDMLQWHTRVGLLVLFLLVFRLCWGIW
                    10
                              20
                                        3.0
                                                 40
                                                           50
                                                                     60
                    70
                              80
                                        90
                                                100
m152.pep
            {\tt GSDTARFSRFVQGWAGIRGYLKNGIPEHIQPGHNPLGALMVVALLAAVSFQVGTGLFAAD}
            g152
            GSDTARFSRFVRGWAGIRGYLKNGIPEHIQPGHNPLGALMVVALLAAVSFQVGTGLFAAN
```

80

90

100

110

BNSDOCID: <WO___9957280A2_J_>

سال د

	130	140	150	160	170	180
m152.pep	ENTFSTNGYLNHL	VSEHTGSLMR:	KIHLNFFKLLA	VFSAIHIAAV <i>I</i>	AYRVFKKKNLI	LPMI
g152	ENTESTNOVI NULL	:				- 111
9132	ENTFSTNGYLNHL	VSEHTGSLIK 140	KIHLNFFKLLA 150			
	200	140	150	160	170	180
	190	200	210	219		
m152.pep	TGFKYIEGKTSIR	FAGKAALAAAI	LSVASLAAAAI	LLLSX		
-150		, , , , , , , , , , ,				
g152	TGFKYIEGKTSIR			LLLSX		
	190	200	210			
The following	g partial DNA se	anence wa	idontified	<b>3</b> 7	:4:3: -CEO	TD (00)
a152.se	2 barnar 13.4V. 20	quence was	s identified i	III IV. mening	mais <seq< td=""><td>ID 623&gt;:</td></seq<>	ID 623>:
	1 ATGAAAAACA	AAACCAAAGT	CTGGGACTTC		· ••••••••••••••••••••••••••••••••••••	
5	1 GCTGCTTGCC	GCATCCCTAC	CCTTTATGTG	GTATAGCGCG	AAAACCGGCG	
10	Ol GCGATATGCT (	GCAATGGCAC	ACGCGCGTCG	GGCTGTTTAT	CCTTTTTCCTC	
15 20		GCCTCTGCTG	GGGCATTTGG	GGCAGCGATA	CCGCCCGTTT	
25		STCCGCGGAT	GGTCGGGTAT	CAGAGAGTAT	ATGAAAAACG	
30	1 GTCGTTGCGC	TTTTGGCCGC	CGTGTCGTTC	' ACCCUTTGGG	CACCCCTGATG	
35	1 TGCCGCCGAT (	STAAACACCT	TCAGCACCAA	CGGCTACCTC	AACCATTTGG	
40	1 TTTCCGAACA T	PACGGGCAGC	CTTATGCGGA	AAATCCATCT	CAACTTTTTC	
45	1 AAACTGCTCG (	CCGTTTTTTC	CGCAGTCCAC	ATCGCCGNCG	TCGCCGCATA	
50 55		AAAAAGAAAA	ACCTCGTCCT	CCCGATGATA	ACCGGCTTCA	
60		PATCGGTTGC	CCCCCTTCCC	TIGCCGGCAA	AGCCGCGCTT	
65	1 GTCCTGA	0001100	0000011000	GCAGCCGCCA	TCCTGCTCCT	
ment +		_				
This correspon	nds to the amino	acid seque	nce <seq i<="" td=""><td>D 624; ORF</td><td>152.a&gt;:</td><td></td></seq>	D 624; ORF	152.a>:	
a152.pe						
5.	1 MKNKTKVWDF I	PTRLFHWLLA	ASLPFMWYSA	KTGGDMLQWH	TR <u>VGLFILFL</u>	
10:		VCTCI FAAD	VRGWSGIREY	MKNGIPEHVQ	PGHNPLGALM	
15:	1 KLLAVFSAVH I	AXVAAYRVF	KKKNI.VI.PMT	TGEKVIEGET	LMRKIHLNEE	
20:	1 AAALSVAALA A	AAILLLS*		TOTKTIEGKT	PIKENGKANE	
150/ 150	0.4.00/.14					
m152/a152	94.0% identity in	n 218 aa ov	erlap			
150			:0 3		50	60
m152.pe	-	LPTRLFHWLL	AASLPFMWYS	AKAGGDMLQWH	PRVGLFVLFLL	VFRLCWGIW
a152	MKNKTKVWD	;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;		:         NKTCCDMIOWU		TED T CHICATA
	1		0 30		50	FRLCWGIW 60
	_					00
m152.pep			0 90	0 100	110	120
misz.pep		FVQGWAGIRG	YLKNGIPEHI(	OPGHNPLGALM	VVALLAAVSFQ	/GTGLFAAD
a152	GSDTARFSR	FVRGWSGIRE	YMKNGTPEHV	IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	 /VALLAAVSFQ	
	7	0 8	0 90		110	120
		_				220
m152.per	13				170	180
m132.pet		PNHTASERIC	PUNKKIHINE	FKLLAVFSAIH:	(AAVAAYRVFKI	KNLILPMI
a152	VNTFSTNGY	LNHLVSEHTG	SLMRKTHLNF	TKI.I.AVFSAVH	I I I I I I I I I I I I I I I I I I I	KNIT VI DMT
	13	0 14	0 150		170	180
m152.per	19		0 210 LAAALSVASLA	219		
manz. per		. D. K. F. H. G. K. A. A. I. I. I. I. I. I. I. I. I. I. I. I. I.	LAAALSVASLA 	AAAAILLLSX		
a152	TGFKYIEGK	TSIRFAGKAA	LAAALSVAAL <i>i</i>	YS:TJ.TTAAA		
	19					

The following partial DNA sequence was identified in N. gonorrhoeae <SEO ID 625>: g153.seg

```
atggggtttg cttaCAgtat gacgtatatc gaggtCGGGa taccggaggc
  1
     ggcatccgtc ctttCgctGC CCGAGATgat gcgcctgatG GTGTTtCagg
     attATGGTTT TrtggcCGAA GTGATGTTTG TGctgaCTTT cGGCGcgcCG
 151 GTTCTGTTLC TGCTGCTGTG CCTGTATGTC TATGCCGCGC TGATACGGAA
 201 ACAGGCGTAT CCTGCGCTGC GTTTGGCAAC GCGTGTGATG GTGCGCTTGA
 251 GGCAGGCGAT GATGGTGGAT GTGTTTTTTG TTTCCACTCT GGTGGCGTAT
 301 ATCAAGCTCT CGTCTGTGGC AAAGGTTCGC TTCGGGCCGG CGTTTTATCT
 351 GATGTTCGCG CTGTCGGTTA TGCTGATTCG GACTTCGGTA TCGGTTCCCC
 401 AGCATTGGGT GTATTTCCAA ATCGGGCGGC TGACGGGGAA TAATGCGGTT
 451 CAGACGGCAT CGGAAGGCAA AACCTGTTGC AGCCGCTGCC TGTATTTccg
 501 cgacAGTgcc gaatccCCCT GCGGGGTGTg cgGCGcggaA CTgtacggcg
 551 gacggccgaa aagtCTGAGt atttCgtCGG CGTTTCTgac ggcggcggTT
     GTTTTGTATT TCCctgCcaa TATCctgccg attaTGAttt cgtccAATCc
 651 tgccgccacg GAGGcCAACA CCATCTTTAG CGGCATCGCT TATATGTGGG
 701 ACGAGGGCGATT GCGGCGGTTA TTTTCAGCGC GAGTATTTTG
 751 GTGCCGGTGC TGAAGATTGC GGCAATGTCG GTTTTGATTG CGGCGGCACG
 801 GTTCGCTTTG CCGGCGGGCG CAAAGAAATT GTCGCACCTC tacCGCATCA
 851 CCGAAGCGGT CGGCCGCTGG TCGATGATTG ATATTTTTGT GATTATTATT
 901 TTGATGTGTT CGTTCCacaC TTATGCCGCG CGCGTCATTC CGGGCAGTGC
     GGCAGTCTAT TTCTGCCTGG TCGTGATTTT GACGATGCTG TCCGCCTATT
1001 ATTTCGACCC GCGCCTGCTT TGGGACAAAC GCGCTTCAGA CGGCATTGCT
1051 TTCAACGAAA CGGAAAAATA TGACTGA
```

This corresponds to the amino acid sequence <SEQ ID 626; ORF 153.ng>: g153.pep

```
MGFAYSMTYI EVGIPEAASV LSLPEMMRLM VFQDYGFLAE VMFVLTFGAP
    VLFLLLCLYV YAALIRKQAY PALRLATRVM VRLRQAMMVD VFFVSTLVAY
 51
101 IKLSSVAKVR FGPAFYLMFA LSVMLIRTSV SVPQHWVYFQ IGRLTGNNAV
151
    QTASEGKTCC SRCLYFRDSA ESPCGVCGAE LYGGRPKSLS ISSAFLTAAV
     VLYFPANILP IMISSNPAAT EANTIFSGIA YMWDEGDRLI AAVIFSASIL
    VPVLKIAAMS VLIAAARFAL PAGAKKLSHL YRITEAVGRW SMIDIFVIII
     LMCSFHTYAA RVIPGSAAVY FCLVVILTML SAYYFDPRLL WDKRASDGIA
301
    FNETEKYD*
```

The following partial DNA sequence was identified in N. meningitidis <SEQ ID 627>: m153.seq

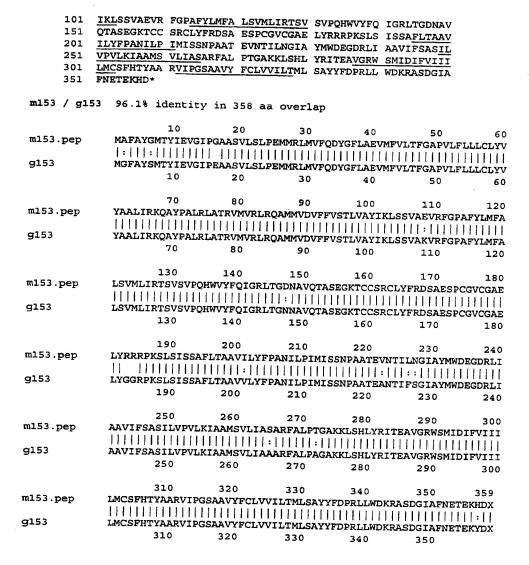
```
ATGGCGTTTG CTTACGGTAT GACGTATATC GAGGTCGGGA TACCGGGTGC
  1
 51 GGCATCCGTC CTTTCGCTGC CCGAGATGAT GCGCCTGATG GTGTTTCAGG
101 ATTATGGTTT TTTGGCCGAA GTGATGTTTG TGCTGACTTT CGGCGCGCCG
151 GTTCTGTTTC TGCTGCTGTG CCTGTATGTC TATGCCGCGC TGATACGGAA
201 ACAGGCGTAT CCTGCGCTGC GTTTGGCAAC GCGTGTGATG GTGCGCTTGA
251 GACAGGCGAT GATGGTGGAT GTGTTTTTTG TTTCCACTTT GGTGGCGTAT
301 ATCAAGCTCT CGTCTGTGGC AGAGGTTCGC TTCGGGCCGG CGTTTTATCT
351 GATGTTCGCG CTGTCAGTTA TGCTGATTCG GACTTCGGTA TCGGTTCCCC
401 AGCATTGGGT GTATTTTCAA ATCGGGCGGC TGACGGGGGA TAATGCGGTT
451 CAGACGGCAT CGGAAGGTAA AACCTGTTGC AGCCGCTGCC TGTATTTCCG
501 CGACAGTGCC GAATCCCCCT GCGGCGTGTG CGGTGCGGAA CTGTACCGCC
551 GACGGCCGAA AAGTCTGAGT ATTTCGTCGG CGTTTCTGAC GGCGGCGGTT
601 ATTTGTATT TCCCTGCCAA TATCCTGCCG ATTATGATTT CGTCCAATCC
651
     TGCCGCCACG GAGGTCAATA CCATCCTTAA CGGCATCGCT TATATGTGGG
701 ACGAGGCGA CAGGCTGATT GCGGCGGTTA TTTTCAGCGC GAGTATTTTG
751 GTGCCGGTAC TGAAGATTGC GGCAATGTCG GTTTTGATTG CGTCCGCCCG
801 CTTCGCTTTG CCAACGGGTG CAAAGAAATT GTCGCACCTC TACCGCATCA
851 CCGAAGCGGT CGGCCGCTGG TCGATGATTG ATATTTTTGT GATTATTATT
901 TTGATGTGTT CGTTCCACAC TTATGCCGCG CGCGTCATTC CGGGCAGTGC
951 GGCAGTCTAT TTCTGCCTGG TCGTGATTCT GACGATGCTG TCCGCCTATT
1001 ATTTCGACCC GCGCCTGCTT TGGGACAAAC GCGCTTCAGA CGGCATTGCT
1051 TTCAATGAAA CGGAAAAACA TGACTGA
```

This corresponds to the amino acid sequence <SEQ ID 628; ORF 153>: m153.pep

- MAFAYGMTYI EVGIPGAASV LSLPEMMRLM VFQDYGFLAE VMFVLTFGAP
- VLFLLLCLYV YAALIRKQAY PALRLATRVM VRLRQAMMVD VFFVSTLVAY

BNSDOCID: <WO___9957280A2_I_>

1.474



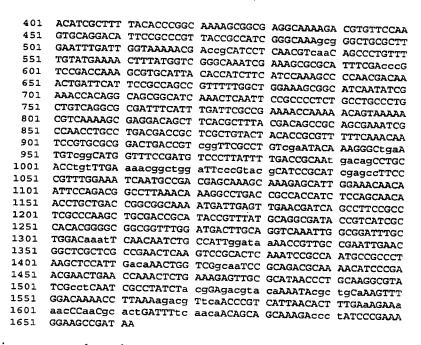
### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 629>:

a153.seq				_	_
1	ATGGCGTTTG	CTTACGGTAT	GACGTATATC	GAGGTCGGGA	TACCGGGTGC
51	GGCATCCGTC	CTTTCGCTGC	CCGAGATGAT	GCGCCTGATG	GTGTTTCAGG
101	ATTATGGTTT	TTTGGCCGAA	GTGATGTTTG	TGCTGACCTT	CGGCGCGCCG
151	GTTCTGTTTC	TGCTGCTGTG	CCTGTATGTC	TATGCCGCGC	TGATACGGAA
201	ACAGGCGTAT	CCTGCGCTGC	GTTTGGCAAC		GTGCGCTTGA
251	GACAGGCGAT	GATGGTGGAT	GTGTTTTTTG	TTTCCACTTT	GGTGGCGTAT
301	ATCAAGCTCT	CGTCTGTGGC		TTCGGATCGG	CGTTTTATCT
351	GATGTTCGCG	CTGTCGGTTA		GACTTCGGTA	
401	AGCATTGGGT	GTATTTTCAA		TGACGGGGGA	TAATGCGGTT
451	CAGACGGCAT	CGGAAGGTAA		AGCCGCTGCC	TGTATTTCCG
501	CGACAGTGCC	GAATCCCCCT	GCGGCGTGTG		CTGTACCGCC
551	GACGGCCGAA	AAGTCTGAGT	ATTTCGTCGG	CGTTTCTGAC	GGCGGCGGTT
601	ATTTTGTATT	TCCCTGCCAA	TATCCTGCCG	ATTATGATTT	CGTCCAATCC
651	TGCCGCCACG	GAGGTCAATA	CCATCCTTAA	CGGCATCGCT	TATATGTGGG
701	ACGAGGGCGA	CAGGCTGATT	GCGGCGGTTA	TTTTCAGCGC	GAGTATTTTG
751	GTGCCGGTAC	TGAAGATTGC		GTTTTGATTG	CGTCCGCCCG
801		CCAACGGGTG			TACCGCATCA
				-+	IACCOCATCA

	901 TT 951 GO 1001 AT	GATGTGTT CAGTCTAT TTCGACCC	CGTTCCACAC	TTATGCCGCG TCGTGATTCT TGGGACAAAC	GACGATGCTG	GATTATTATT CGGGCAGTGC TCCGCCTATT CGGCATTGCT	
	esponds to	the amin	o acid seque	ence <seq i<="" td=""><td>D 630; ORF</td><td>153.a&gt;:</td><td></td></seq>	D 630; ORF	153.a>:	
ar	1 MA 51 VI 101 IK 151 QT 201 II 251 VE 301 LM	FLLLCLYV LSSVAEVR ASEGKTCC YFPANILP VLKIAAMS	YAALIRKQAY FGSAFYLMFA SRCLYFRDSA IMISSNPAAT VLIASARFAL	PALRLATRVM LSVMLIRTSV ESPCGVCGAE EVNTILNGIA PTGAKKLSHL	VRLRQAMMVD SVPQHWVYFQ LYRRRPKSLS YMWDEGDRLI YRITEAVGRW	VMFVLTFGAP VFFVSTLVAY IGRLTGDNAV ISSAFLTAAV AAVIFSASIL SMIDIFVIII WDKRASDGIA	
m153/a1:	53 99.7%	6 identity	in 358 aa ov	-	_		
m15	3. <b>pe</b> p	MAFAYGM	YIEVGIPGAA	SVLSLPEMMRL	0 40 MVFQDYGFLAE	50 VMFVLTFGAPVI 	60 E <b>FLLL</b> CLYV
a15	3	MAFAYGM	YIEVGIPGAA	SVLSLPEMMRL	MVFQDYGFLAE 0 40	VMFVLTFGAPVI	FLLLCLYV 60
4-	_				0 100		120
m15 a15	3.pep		1111111111	1111111111	11111111111	IKLSSVAEVRFO             IKLSSVAEVRFO	1111111
					0 100		120
m15	3.pep	LSVMLIR	SVSVPQHWVY	40 15 FOIGRLTGDNA	VQTASEGKTCC	170 SRCLYFRDSAES	180 SPCGVCGAE
a15	3	LSVMLIRT	SVSVPQHWVY	FQIGRLTGDNA 40 15	VQTASEGKTCC		PCGVCGAE 180
m15	3.pep	LYRRRPKS	LSISSAFLTA	00 21 AVILYFPANIL	PIMISSNPAAT	230 EVNTILNGIAYN	240 WDEGDRLI
a15	3	LYRRRPKS	LSISSAFLTA	AVILYFPANIL	PIMISSNPAAT		WDEGDRLI 240
<b>m1</b> 5	3.pep	AAVIFSAS	ILVPVLKIAA	60 27 MSVLIASARFA	LPTGAKKLSHL	290 YRITEAVGRWSM	300 IIDIFVIII
a15	3	AAVIFSAS	ILVPVLKIAA	 MSVLIASARFA 50 27	LPTGAKKLSHL		IIIIIIIII IIDIFVIII 300
m15	3.pep	LMCSFHTY	'AARVIPGSAA'	20 33 VYFCLVVILTM	LSAYYFDPRLL	WDKRASDGIAFN	359 ETEKHDX
a15	3	LMCSFHTY	 AARVIPGSAA		 LSAYYFDPRLL		111111
The follo	wing parti	al DNA so	equence was	identified in	n <i>N. gonorrh</i>	oeae <seq 1<="" td=""><td>D 631&gt;:</td></seq>	D 631>:
101 151 201 251 301	CAAAACAAC TGATTGCCGC GTGGTTACGC GGTCATTAAC TGCGCGACGA GTATCCGGCC	C ACCTTCCT CGGCTGGC CTCTTGATG CGAAAAAG CCCAAAAAG CTCATCCGC	CT CCGCCGTCT TT TGGGTTAAG GA CAGCGCGGA CA TCGATGTCG GC GTGGAAGTT AG CGATACCCA	CACGCTCAAG G GCTGGTCCCG AAATCCGCAA GGCATCGAAG G ACGCGTTACC A CTGCCCAACT G TTTTGGGTGG G TACGCTGCTT	CTGATCGCGC CAGGGGGCCT TCAACAATAC CGAATCAAAC CAATGCGGAC TCAAGCCGCG		

BNSDOCID: <WO___9957280A2_J_>

المعالم المعالم المعالم المعالم



## This corresponds to the amino acid sequence <SEQ ID 632; ORF 154.ng>:

```
MTDNSPPPNG HAQARVRKNN TFLSAVWLVP LIALIAGGWL WVKEIRNRGP
     VVTLLMDSAE GIEVNNTVIK VLSIDVGRVT RIKLRDDQKG VEVTAQLNAD
 51
     VSGLIRSDTQ FWVVKPRIDQ SGVTGLGTLL SGSYIAFTPG KSGEAKDVFQ
101
     VQDIPPVTAI GQSGLRLNLI GKNDRILNVN SPVLYENFMV GQIESAHFDP
201
     SDQSVHYTIF IQSPNDKLIH SASRFWLESG INIETTGSGI KLNSAPLPAL
     LSGAISFDSP KTKNSKNVKS EDSFTLYDSR SEIANLPDDR SLYYTAFFKQ
251
301 SVRGLTVGSP VEYKGLNVGM VSDVPYFDRN DSLHLFENGW IPVRIRIEPS
351 RLEINADEQS KEHWKQQFQT ALNKGLTATI SSNNLLTGGK MIELNDQPSA
     SPKLRPHTVY AGDTVIATRG GGLDDLQVKL ADLLDKFNNL PLDKTVAELN
401
    GSLAELKSAL KSANAALSSI DKLVGNPQTQ NIPNELNQTL KELRITLQGV
451
     SPQSPIYGDV QNTLQSLDKT LKDVQPVINT LKEKPNALIF NNSSKDPIPK
501
```

## The following partial DNA sequence was identified in N. meningitidis <SEQ ID 633>: m154.seq

1	ATGACTGACA	ACAGCCCTCC	TCCAAACGGA	CACGCCCAAG	CACGCGTCCG
51	CAAAAACAAC	ACCTTCCTCT	CTGCCGTCTG		CTGATCGCGC
101	TGATTGCCGG				CAGGGGGCCT
151	GTGGTTACGC	TCTTGATGGA	CAGCGCGGAA		TCAACAATAC
201	GGTCATCAAA	GTATTGAGCA	TCGATGTCGG		CGAATCAAAC
251		CCAAAAAGGC			CAATGCGGAC
301	GTATCCGGCC	TCATCCGCAG	CGATACCCAG		
351			CCGGTTTGGG		TCAAGCCGCG
401	ACATCGCCTT	TACACCCGGC	AAAAGCGACG	ACCCARAGE	TCGGGTTCGT
451	GTGCAGGACA	TTCCGCCCGT	TACCGCCATC		
501		GGTAAAAACG			GGCTGCGCTT
551		TTTTATGGTC		CAACGTCAAC	AGCCCTGTTT
601		GCGTGCATTA		AAAGCGCGCA	TTTCGACCCG
651	ACTGATTCAT			ATCCAAAGCC	CCAACGACAA
701				GGAAAGCGGC	ATCAATATCG
751	CTGTCGGGCG	CAGCGGCATC	AAACTCAATT	CCGCCCCTCT	GCCTGCCCTG
801					ACAGTAAAAA
851		GAAGACAGCT			AGCGAAGTCG
901	CCAACCTGCC			ACACCGCGTT	TTTCAAACAA
	TCCGTGCGCG			GTCGAGTACA	AAGGGCTGAA
951		GTTTCCGACG		CGACCGCAAC	GACAGCCTGC
1001	ACCTGTTTGA		ATACCCGTAC	GCATCCGCAT	TCAACCTTCC
1051	CGTTTGGAAA	TCAATGCCGA	CGAACAAAGC	AAAGAACATT	GGAAACAACA
1101	ATTTCAGACG	GCCTTAAACA	AAGGCCTGAC	CGCCACCATC	TCCACCAACA
1151	ACCTGCTGAC	CGGAAGCAAA	ATGATTGAGT	TGAACGATCA	GCCTTCCGCA

1201 1251 1301 1351 1401 1451 1501 1551 1601 1651	TCACCTAAGC TGCGACCGCA TACCGTTTAT GCAGGCGATA CCGTTATCGC GACCCAGGCC GGCGGTTTGG ACGATTTGCA GGTCAAATTG GCGGATTTGC TGGACAAGTT CGACAAACTG CCTTTAGATA AGACGGTTGC CGAATTGAAC GGTTCGCTTG CCGAGGCTCAA ATCCACACTC AAATCTGCCA ATGCCGCCCT AAGCTCCATC GACAAACTGG TCGGCAAACC GCAGACACAA AACATTCCGA ACGAACTGAA CCCAAACCTG AAAGAGTTGC GCACAACCCT GCAAAGCGTA TCGCCGCAAT CGCCTATCTA CGGCGACGTA CAAAATACGC TGCAAAGTTT GGACAAAACT TTAAAAGACG TTCAACCCGT GATTAATACT TTGAAAGAAA AACCCAACGC GCTGATTTTC AACAGCAGCA GCAAAGACCC TATCCCGAAA GGAAGCCGAT AA
	responds to the amino acid sequence <seq 154.a="" 634="" id="" orf="">:</seq>
m154.pep	MTDNSPPPNG HAQARVRKNN TFLSAVWLVP LIALIAGGWL WVKEIRNRGP
51	VVTLLMDSAE GIEVNNTVIK VLSIDVGRVT RIKLRDDQKG VEVTAQLNAD
101 151	VSGLIRSDTQ FWVVKPRIDQ SGVTGLGTLL SGSYIAFTPG KSDEAKDVFQ VQDIPPVTAI GQSGLRLNLI GKNDRILNVN SPVLYENFMV GQVESAHFDP
201	SDQSVHYTIF IQSPNDKLIH SASRFWLESG INIETTGSGI KLNSAPLPAL
251	LSGAISFDSP KTKNSKNVKS EDSFTLYDSR SEVANLPDDR SLYYTAFFKO
301	SVRGLTVGSP VEYKGLNVGV VSDVPYFDRN DSLHLFENGW IPVRIRIEPS
351 401	RLEINADEQS KEHWKQQFQT ALNKGLTATI SSNNLLTGSK MIELNDQPSA SPKLRPHTVY AGDTVIATQG GGLDDLQVKL ADLLDKFDKL PLDKTVAELN
451	GSLAELKSTL KSANAALSSI DKLVGKPQTQ NIPNELNOTL KELRTTLOGV
501	SPQSPIYGDV QNTLQSLDKT LKDVQPVINT LKEKPNALIF NSSSKDPIPK
551	GSR*
m154 / g1	97.8% identity in 553 aa overlap
	10 20 30 40 50 60
m154.pep	MTDNSPPPNGHAQARVRKNNTFLSAVWLVPLIALIAGGWLWVKEIRNRGPVVTLLMDSAE 
g154	MTDNSPPPNGHAQARVRKNNTFLSAVWLVPLIALIAGGWLWVKEIRNRGPVVTLLMDSAE
	10 20 30 40 50 60
	70 80 90 100 110 120
m154.pep	GIEVNNTVIKVLSIDVGRVTRIKLRDDQKGVEVTAQLNADVSGLIRSDTOFWVVKPRIDO
~154	
g154	GIEVNNTVIKVLSIDVGRVTRIKLRDDQKGVEVTAQLNADVSGLIRSDTQFWVVKPRIDQ 70 80 90 100 110 120
	120
m154.pep	130 140 150 160 170 180 SCUTTCL CTLL SCSVIA ETRICAGREDA APPRICA CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CONTROL TO CO
mist.pep	SGVTGLGTLLSGSYIAFTPGKSDEAKDVFQVQDIPPVTAIGQSGLRLNLIGKNDRILNVN 
g154	SGVTGLGTLLSGSYIAFTPGKSGEAKDVFQVQDIPPVTAIGQSGLRLNLIGKNDRILNVN
	130 140 150 160 170 180
m154.pep	190 200 210 220 230 240 SPVLYENFMVGQVESAHFDPSDQSVHYTIFIQSPNDKLIHSASRFWLESGINIETTGSGI
g154	SPVLYENFMVGQIESAHFDPSDQSVHYTIFIQSPNDKLIHSASRFWLESGINIETTGSGI
	190 200 210 220 230 240
	250 260 270 280 290 300
m154.pep	KLNSAPLPALLSGAISFDSPKTKNSKNVKSEDSFTLYDSRSEVANLPDDRSLYYTAFFKQ 
g154	KLNSAPLPALLSGAISFDSPKTKNSKNVKSEDSFTLYDSRSEIANLPDDRSLYYTAFFKQ
	250 260 270 280 290 300
	310 320 330 340 350 360
m154.pep	SVRGLTVGSPVEYKGLNVGVVSDVPYFDRNDSLHLFENGWIPVRIRIEPSRLEINADEOS
g154	
3	310 320 330 340 350 360
	370 380 390 400 410 420
m154.pep	KEHWKQQFQTALNKGLTATISSNNLLTGSKMIELNDQPSASPKLRPHTVYAGDTVIATOG
-154	
g154	KEHWKQQFQTALNKGLTATISSNNLLTGGKMIELNDQPSASPKLRPHTVYAGDTVIATRG

BNSDOCID: <WO___9957280A2_I_>

	370	380	390	400	410	420
m154.pep	430 GGLDDLQVKLADLLD            GGLDDLQVKLADLLD 430	11::11111		1111:1111	11111111	11.1111
m154.pep	490 NIPNELNQTLKELRT            NIPNELNQTLKELRI 490	1111111	111111111		111111111	11111
m154.pep	550 NSSSKDPIPKGSRX  :          NNSSKDPIPKGSRX 550					

## The following partial DNA sequence was identified in N. meningitidis <SEQ ID 635>:

```
ATGACTGACA ACAGCCCTCC TCCAAACGGA CACGCCCAAG CACGCGTCCG
       CAAAAACAAC ACCTTCCTCT CTGCCGTCTG GCTGGTTCCG CTGATCGCGC
   51
       TGATTGCCGG CGGCTGGCTT TGGGTTAAGG AAATCCGCAA CAGGGGGCCT
  101
       GTGGTTACGC TCTTGATGGA CAGCGCGGAA GGCATTGAGG TCAACAATAC
  151
      GGTCATCAAA GTATTGAGCA TCGATGTCGG ACGCGTTACC CGAATCAAAC
 201
      TGCGCGACGA CCAAAAAGGC GTGGAAGTAA CCGCCCAACT CAATGCGGAC
 251
      GTATCCGGCC TCATCCGCAG CGATACCCAG TTTTGGGTGG TCAAGCCGCG
 301
      TATCGACCAA AGCGGCGTAA CCGGTTTGGG TACGCTGCTT TCGGGTTCGT
 351
      ACATCGCCTT TACACCCGGC AAAAGCGACG AGGCAAAAGA CGTGTTCCAA
 401
      GTGCAGGACA TTCCGCCCGT TACCGCCATC GGGCAAAGCG GGCTGCGCTT
 451
      GAATTTGATT GGTAAAAACG ACCGCATCCT CAACGTCAAC AGCCCTGTTT
 501
      TGTATGAAAA CTTTATGGTC GGGCAAGTCG AAAGCGCGCA TTTCGACCCG
 551
      TCCGACCAAA GCGTGCATTA CACCATCTTC ATCCAAAGCC CCAACGACAA
 601
      ACTGATTCAT TCCGCCAGCC GTTTCTGGCT GGAAAGCGGC ATCAATATCG
 651
      AAACCACAGG CAGCGGCATC AAACTCAATT CCGCCCCTCT GCCTGCCCTG
 701
      CTGTCGGGCG CGATTTCATT TGATTCGCCG AAAACCAAAA ACAGTAAAAA
 751
      CGTCAAAAGC GAAGACAGCT TCACGCTTTA CGACAGCCGC AGCGAAGTCG
 801
      CCAACCTGCC TGATGACCGT TCGCTGTACT ACACCGCGTT TTTCAAACAA
 851
      TCCGTGCGCG GACTGACCGT CGGTTCGCCT GTCGAGTACA AAGGGCTGAA
 901
      TGTCGGCGTG GTTTCCGATG TTCCTTATTT CGACCGCAAC GACAGCCTGC
 951
1001 ACCTGTTTGA AAACGGCTGG ATTCCCGTAC GCATCCGTAT TGAGCCTTCC
      CGTTTGGAAA TCAATGCCGA CGAACAAGC AAAGAACATT GGAAACAACA
1051
     ATTTCAGACG GCCTTAAACA AAGGCCTGAC CGCCACCATC TCCAGCAACA
1101
      ACCTGCTGAC CGGCAGCAAA ATGATTGAGT TGAACGATCA GCCTTCCGCC
1151
      TCGCCCAAGC TGCGACCGCA TACCGTTTAT GCAGGCGATA CCGTTATCGC
1201
     GACCCAGGGC GGCGGTTTGG ACGATTTGCA GGTCAAATTG GCGGATTTGC
     TGGACAAGTT CGACAAACTG CCTTTAGATA AGACGGTTGC CGAATTGAAC
1301
     GGTTCGCTTG CCGAGCTCAA ATCCACACTC AAATCTGCCA ATGCCGCCCT
1351
     AAGCTCCATC GACAAACTGG TCGGCAAACC GCAGACACAA AACATTCCGA
1401
     ACGAACTGAA CCAAACCCTG AAAGAGTTGC GCACAACCCT GCAAGGCGTA
1451
     TCGCCTCAAT CGCCTATCTA CGGCGACGTA CAAAATACGC TGCAAAGTTT
1501
     GGACAAAACC TTAAAAGACG TTCAACCCGT CATTAACACT TTGAAAGAAA
1551
1601
     AACCCAACGC GCTGATTTTC AACAGCAGCA GCAAAGACCC TATCCCGAAA
1651
     GGAAGCCGAT AA
```

### This corresponds to the amino acid sequence <SEQ ID 636; ORF 154.a>:

1 MTDNSPPPNG HAQARVRKNN T<u>FLSAVWLVP</u> LIALIAGGWL WVKEIRNRGP 51 VVTLLMDSAE GIEVNNTVIK VLSIDVGRVT RIKLRDDQKG VEVTAQLNAD 101 VSGLIRSDTQ FWVVKPRIDQ SGVTGLGTLL SGSYIAFTPG KSDEAKDVFQ

a154

a154

a154

m154.pep

m154.pep

370

430

490

490

380

440

440

500

500

151 201 251 301 351 401 451 501	VQDIPPVTAI GQSGLRLNI SDQSVHYTIF IQSPNDKLI LSGAISFDSP KTKNSKNVF SVRGLTVGSP VEYKGLNVG RLEINADEQS KEHWKQQFC SPKLRPHTVY AGDTVIATO GSLAELKSTL KSANAALSS SPQSPIYGDV QNTLQSLDF GSR*	IH SASRFWLESG KS EDSFTLYDSR GV VSDVPYFDRN QT ALNKGLTATI QG GGLDDLQVKL SI DKLVGKPOTO	INIETTGSGI SEVANLPDDR DSLHLFENGW SSNNLLTGSK ADLLDKFDKL NIPNELNOTL	KLNSAPLPAL SLYYTAFFKQ IPVRIRIEPS MIELNDQPSA PLDKTVAELN KELRTTLOGV	
m154/a154 10	0.0% identity in 553 aa	overlap			
m154.pep	10 MTDNSPPPNGHAQARVR	20 30 RKNNTFLSAVWLV	PLIALIAGGWLW	50 VKEIRNRGPVVTI	60 LLMDSAE
a154		RKNNTFLSAVWLV 20 30	PLIALIAGGWLW	!	LINDSAE 60
	70	00			
m154.pep	70 GIEVNNTVIKVLSIDVG	80 90 GRVTRIKLRDDQKO	GVEVTAQLNADV	110 SGLIRSDTQFWVV	120 /KPRIDQ
a154	GIEVNNTVIKVLSIDVG 70	RVTRIKLRDDQKO	SVEVTAQLNADV	'SGLIRSDTQFWV\ 110	/KPRIDQ 120
-154 ··		140 150		170	180
m154.pep	SGVTGLGTLLSGSYIAF	TPGKSDEAKDVF(	QVQDIPPVTAIG	QSGLRLNLIGKNI	RILNVN
a154	SGVTGLGTLLSGSYIAF	TPGKSDEAKDVF	QVQDIPPVTAIG	QSGLRLNLIGKNE	RILNVN
	130	140 150	160	170	180
m154.pep	SPVLYENFMVGQVESAH	200 210 FDPSDQSVHYTI	TIOSPNDKLIHS	230 ASRFWLESGINIE	240 TTGSGI
a154				1111111111111	111111
		200 210	220	230	240
m154.pep		260 270		290	300
mro4.pep	KLNSAPLPALLSGAISF		11111111111	111111111111	TELLI
a154	KLNSAPLPALLSGAISF	DSPKTKNSKNVKS 260 270	EDSFTLYDSRS	EVANLPDDRSLYY 290	TAFFKQ .
	310	320 330	240	250	
m154.pep	SVRGLTVGSPVEYKGLN	VGVVSDVPYFDRN	DSLHLFENGWI	350 PVRIRIEPSRLEI	360 NADEQS
a154				PURTRIED SDIET	MADEOS
	310	320 330		350	360
		380 390		410	420
m154.pep	KEHWKQQFQTALNKGLT	ATISSNNLLTGSK	MIELNDOPSAS	PKLRPHTVYAGDT	VIATQG
a154		1111111111111 ATTSSNNTTTCSV			

 ${\tt KEHWKQQFQTALNKGLTATISSNNLLTGSKMIELNDQPSASPKLRPHTVYAGDTVIATQG}$ 

 ${\tt GGLDDLQVKLADLLDKFDKLPLDKTVAELNGSLAELKSTLKSANAALSSIDKLVGKPQTQ}$ 

GGLDDLQVKLADLLDKFDKLPLDKTVAELNGSLAELKSTLKSANAALSSIDKLVGKPQTQ

NIPNELNQTLKELRTTLQGVSPQSPIYGDVQNTLQSLDKTLKDVQPVINTLKEKPNALIF

NIPNELNQTLKELRTTLQGVSPQSPIYGDVQNTLQSLDKTLKDVQPVINTLKEKPNALIF

400

460

460

520

520

410

530

530

420

480

390

450

450

510

510

550
m154.pep NSSSKDPIPKGSRX
|||||||||||||
a154 NSSSKDPIPKGSRX
550

## The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 637>:

```
atGAAaatcg GtatcCCACG CGAGTCAtta tcCGGCGAAA cccgcgtagc
      ctgcAcgccc gCCACCGTTG CCctgctggg caAactAGGC TTTGAAACCG
   51
      TTGtcgaAAG CGGTGCAggt TTGGCGGCAA GTTTggaCGA TGCCGCTTAC
      CAAACAGCAG GCGCAACCGT TGCCGACAAA GCGGCGGTTT GGGCCTGCCC
      TTTAATTTAT AAGGTCAACG CGCCGTCCGA AGGCGAGCTG CCGCTGCTCA
      AAGAAGGTCA AACCATCGTC AGCTTCCTGT GGCCGCGCCA AAACGAGGCT
      TTGGTCGAGG CCTTGCGCGC CAAGAAAGTC AACGCGCTGG CGATGGACAT
 301
      GGTTCCCCGC ATTTCCCGCG CTCAGGCCTT GGACGCTTTG TCTTCAATGG
 351
      CAAACATCAG CGGCTACCGC GCCGTGATTG AAGCCGCCAA CGCCTTCGGC
 401
      CGTTTCTTCA CCGGTCAAAT CACTGCCGCC GGCAAAGTGC CGCCTGCGCA
 451
      GGTTTTGGTG ATTGGCGCCG GTGTGGCGGG TTTGGCGGCA ATCGGTACGG
      CAAATTCGCT CGGCGCAGTG GTGCGCGCGT TCGATACCCG CTTGGAAGTG
 551
      GCGGAACAAA TCGAATCGAT GGGCGGTAAG TTcctGAAAC TCGACTTCCT
      GCAAGAATCG GGCGCCAGCG GAGACGGCLA CGCCAAAGTG ATGAGCGACG
      AATTTATCGC CGCCGAAATG AAGCTCTTTG CCGAACAGGC GAAAGAAGTG
 701
      GACATCATCA TCACCACCGC CGCCATTCCG GGCAAACCCG CTCCCAAGCT
      GATTACCAAA GAAATGGTGG AAAGCATGAA ATCCGGATCC GTCATCGTCG
 801
 851
      ATTTGGCGGC GACGGGCGC AACTGCGAAC TCACCCGACC GGGCGAATTG
      TCCGTAACCG GCAACGGCGT GAAAATCATC GGCTACACCG ACATGGCAAA
 901
 951 CCGCCTTGCC GGACAGTCTT CCCAGCTTTA CGCCACCAAC TTGGTGAACC
1001
      TGACCAAGCT GTTAAGCCCG AACAAAGAcg gcgaAATCAC GCTGGACTTC
      GAAGacgtGA TTATCCGCAA TATGACCGTT ACCCGcgacg gcgaaATCAC
1051
     CTTCCCGCCT CCGccgaTTc aggtTTCcgc ccggccgCAG CAAAcgccgt
1101
     ctgaAAAagc cgcGCCTGCC GCCAagcccg AgccGaaacc tgttCCcctg
     tggaAAAaac tcgCGCCCGC CGCcatcgCC GCCGTATTGG tgctgtgGgt
1201
1251
     cggCgcggtc gcacccgcag CATTCTTGAA CCACTTTATC GTCTTCGTCC
     TCGCCTGCGT CATCGGCTAC CATGTCGTTT GgaacgTCAG CCACTCGCTG
1301
     CACACACCGC TGAtgtcggt aaccaaCgcc atctccGGCA tcatggtcgt
1351
     CGGCGCTG CTGCAAATCG GTCAGGGCaa CGGCttcgtT TCgctGCTGT
     CGTTTGTTGC CATCCTGATT GCCGGCATCA ATATCTTCGG CGGCTTTGCG
1501 GTTACACGGC GTATGCTGAA TATGTTTAAG AAAGGGTAA
```

#### This corresponds to the amino acid sequence <SEQ ID 638; ORF 155.ng>:

```
MKIGIPRESL SGETRVACTP ATVALLGKLG FETVVESGAG LAASLDDAAY

OTAGATVADK AAVWACPLIY KVNAPSEGEL PLLKEGOTIV SFLWPRQNEA

LVEALRAKKV NALAMDMVPR ISRAQALDAL SSMANISGYR AVIEAANAFG

RFFTGOITAA GKVPPAQVLV IGAGVAGLAA IGTANSLGAV VRAFDTRLEV

AEQIESMGGK FLKLDFLQES GGSGDGYAKV MSDEFIAAEM KLFAEQAKEV

DIIITTAAIP GKPAPKLITK EMVESMKSGS VIVDLAATGG NCELTRPGEL

SVTGNGVKII GYTDMANRLA GQSSQLYATN LVNLTKLLSP NKDGEITLDF

SUTGNGVKII TRDGEITFPP PPIQVSARPQ QTPSEKAAPA AKPEPKPVPL

401 WKKLAPAAIA AVLVLWVGAV APAAFLNHFI VFVLACVIGY HVVWNVSHSL

451 HTPLMSVTNA ISGIMVVGAL LQIGQGNGFV SLLSFVAILI AGINIFGGFA
```

## The following partial DNA sequence was identified in N. meningitidis <SEQ ID 639>:

```
1 ATGAAAATCG GTATCCCACG CGAGTCATTA TCCGGCGAAA CCCGCGTCGC
51 CTGTACGCCC GCCACCGTCG CCCTGCTGGG CAAACTGGGC TTTGAAACCG
101 TTGTCGAAAG CGGTGCAGGT TTGGCGGCAA GTTTGGACGA TGCCGCTTAC
151 CAAACAGCAG GCGCAACCGT TGCCGACAAA GCGGCGGTTT GGGTCTGCCC
201 TTTGATTTAT AAGGTCAACG CGCCGTCCGA ACAGGAACTG CCGCTTTTGA
251 ACGAAGGTCA AACCATCGTC AGCTTCCTGT GGCCGCGCCA AAACGAGGCT
```

301	TTGGTCGAAG	CCTTGCGCGC	CAAGAAAGTG	AACGCGCTGG	CGATGGATAT
351	GGTGCCCCGC	ATTTCGCGCG	CGCAGGCTTT	GGACGCTTTG	TCTTCGATGG
401	CAAACATCAG	CGGCTACCGC	GCCGTAATTG	AAGCCGCCAA	CGCCTTCGGC
451	CGTTTCTTCA	CCGGTCAAAT	TACCGCCGCC	GGCAAAGTGC	CGCCCGCGCA
501	GGTTTTGGTG	ATTGGTGCAG	GTGTGGCAGG	TTTGGCGGCG	ATCGGTACGG
551	CAAACTCGCT	CGGCGCAGTG	GTACGCGCGT	TCGATACCCG	CTTGGAAGTG
601	GCGGAACAAA	TCGAATCGAT	GGGCGGCAAG	TTCCTGAAAC	TCGACTTCCC
651	ACAAGAATCG	GGCGGCAGCG	GAGACGGCTA	CGCCAAAGTG	ATGAGCGACG
701	AATTTATCGC	AGCCGAGATG	AAGCTCTTTG	CCGAGCAGGC	GAAAGAAGTG
751	GACATCATCA	TCACCACCGC	CGCCATTCCG	GGCAAACCCG	CGCCCAAGCT
801	GATTACCAAA	GAAATGGTGG	AAAGCATGAA	ATCCGGCTCC	GTCATCGTCG
851	ATTTGGCGGC	GGCGACGGGC	GGCAACTGCG	AACTCACCCG	CCCGGGCGAA
901	TTGTCCGTAA	CCGGCAACGG	CGTGAAAATC	ATCGGCTACA	CCGACATGGC
951	AAACCGCCTT	GCCGGACAGT	CTTCCCAGCT	TTACGCCACC	AACTTGGTCA
1001	ACCTGACCAA	GCTGTTAAGC		ACGGCGAAAT	CACGTTGGAC
1051	TTCGAAGACG	TGATTATCCG	CAACATGACC	GTTACCCACG	ACGGCGAAAT
1101	CACCTTCCCG	CCTCCGCCGA	TTCAAGTTTC	CGCCCAGCCG	CAGCAAACGC
1151	CGTCTGAAAA	AGCCGTGCCT	GCCGCCAAGC	CCGAGCCAAA	ACCCGTTCCC
1201	CTGTGGAAAA	AACTCGCGCC	CGCCGTCATC	GCCGCCGTCT	TGGTACTGTG
1251	GGTCGGCGCG	GTCGCACCCG	CAGCATTCCT	GAACCACTTT	ATCGTGTTCG
1301	TTCTCGCCTG	CGTCATCGGC	TACTACGTCG	TCTGGAACGT	CAGCCACTCG
1351	CTGCACACAC	CGCTGATGTC	GGTAACCAAC	GCCATCTCCG	GCATCATCGT
1401	CGTCGGCGCG	CTGCTGCAAA	TCGGTCAGGG	CAACGGCTTC	GTTTCGCTGC
1451	TGTCGTTTGT	TGCCATCCTG	ATTGCCGGCA	TCAACATCTT	CGGCGGCTTT
1501	GCGGTAACAC	GGCGTATGCT	GAATATGTTT	AAGAAAGGGT	AA

#### This corresponds to the amino acid sequence <SEQ ID 640; ORF 155>:

				,	
m155.pep					
1	MKIGIPRESL	SGETRVACTP	ATVALLGKLG	FETVVESGAG	LAASLDDAAY
51	QTAGATVADK	<b>AAVWVCPLIY</b>	KVNAPSEQEL	PLLNEGQTIV	SFLWPRQNEA
101	LVEALRAKKV	NALAMDMVPR	ISRAQALDAL	SSMANISGYR	AVIEAANAFG
151			IGAGVAGLAA		
201	AEQIESMGGK	FLKLDFPQES	GGSGDGYAKV	MSDEFIAAEM	KLFAEQAKEV
251	DIIITTAAIP	GKPAPKLITK	<b>EMVESMKSGS</b>	VIVDLAAATG	GNCELTRPGE
301	LSVTGNGVKI	IGYTDMANRL	AGQSSQLYAT	NLVNLTKLLS	PNKDGEITLD
351	FEDVIIRNMT	VTHDGEITFP	PPPIQVSAQP	QQTPSEKAVP	AAKPEPKPVP
401	LWKKLAPAVI	AAVLVLWVGA	VAPAAFLNHF	IVFVLACVIG	YYVVWNVSHS
451	LHTPLMSVTN	AISGIIVVGA	LLQIGQGNGF	VSLLSFVAIL	IAGINIFGGF
501	AUTRRMINME				

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from N. gonorrhoeae

ORF 155 shows 97.9% identity over a 513 aa overlap with a predicted ORF (ORF 155.ng) from N. gonorrhoeae:

m155 / g155	97.9% identity	in 513 aa	overlap			
	10	20	30	40	50	60
m155.pep	MKIGIPRESLSGET	RVACTPATVA	LLGKLGFET	/VESGAGLAAS:	LDDAAYQTAG	ATVADK
	1111111111111111	1111111111	1111111111		1111111111	111111
g155	MKIGIPRESLSGET	RVACTPATVA	LLGKLGFET	/VESGAGLAAS:	LDDAAYQTAC	ATVADK
	10	20	30	40	50	60
	70	80	90	100	110	120
m155.pep	<b>AAVW</b> VCPLIYKVNA	PSEQELPLLN	<b>EGQTIVSFL</b>	VPRQNEALVEA:	LRAKKVNALA	MDMVPR
		111 11111:				111111
g155	<b>AAVW</b> ACPLIYKVNA	PSEGELPLL	ŒGQTIVSFLV	PRQNEALVEA:	LRAKKVNAL <i>i</i>	MDMVPR
	70	80	90	100	110	120
	100					
	130	140	150	160	170	180
m155.pep	ISRAQALDALSSMA	NISGYRAVIE	CAANAFGRFFT	GQITAAGKVP	PAQVLVIGAG	VAGLAA
		111111111	1111111111			111111
g155	ISRAQALDALSSMA	NISGYRAVIE	AANAFGRFFT	GQITAAGKVP	PAQVLVIGAG	VAGLAA

			130	140	150	160	170	180
	m155.pep	IGTAN	190 SLGAVVRA	200 FDTRLEVAEQI	210 ESMGGKFLK	220	230	240
	g155	1111	, , , , , , , , ,		1			
			130	200	210	220	230	240
	m155.pep	KLFAE	250 QAKEVDII:	260 ITTAAIPGKPA	270 PKLITKEMVE	280 ESMKSGSVIVD	290	300
	g155		QAKEVDII	[TTAAIPGKPA]				
			250	260	270	280	290	LINIGE
	m155.pep	LSVTG	310 NGVKIIGYT	320 DMANRLAGOSS	330 SQLYATNLVN	340 LTKLLSPNKD	350 GEITLDFEDV	360
	g155		GVKIIGYT	'DMANRLAGOSS	GUYATNLVN			
		300	310	320	330	340	350	11111171
	m155.pep	VTHDGE	370 ITFPPPPI	380 QVSAQPQQTPS	390 EKAVPAAKP	400 EPKPVPLWKKI	410 APAVIAAVL	420 VLWVGA
	<b>g15</b> 5		ITFPPPPI	QVSARPQQTPS	EKAAPAAKP			
		360	370 430	380	390	400	410	
	m155.pep	VAPAAF	LNHFIVFV	440 LACVIGYYVVW	450 NVSHSLHTP	460 LMSVTNAISGI	470 IVVGALLQI	480 GOGNGF
	g155			LACVIGYHVVW	(          NVSHSLHTPI	<b></b>		
		420	430	440	450	460	470	
1	m155.pep	VSLLSF	VAILIAGIN	500 NIFGGFAVTRRI	510 MLNMFKKGX			
•	g155	VSLLSF	VAILIAGIN 490		1LNMFKKGX			
			400	300	510			

# The following partial DNA sequence was identified in N. meningitidis <SEQ ID 641>: a155.seq

1	ATGAAAATC	GTATCCCACG	TGAGTCATTA	TCCGGCGAAA	CCCGCGTCGC
51	CTGTACGCCC	GCCACCGTCG			
101	TTGTCGAAAG	CGGCGCAGGT			
151	CAAGCAGCAG				
201	TTTAATTTAT				
251	AAGAAGGACA				
301	TTGGTCGAAC	011001110010			
351	GGTGCCCCGC	0011000000			
401	CAAACATCAG	***********			
451	CGTTTNTTCA				CGCCTTCGGC
501	GGTTTTGGTG	1100000 HEAT	-11-01-0000CM		CGCCCGCGCA
551			GTGTGGCAGG		ATCGGTACGG
601	CAAACTCGCT		GTACGCGTGT	TCGATACCCG	CCTG. AAGTG
651	GCGGAACAAT	0010121100111	GGGCGGCAAG	TTCCTGAAAC	TCGACTTCCC
	GCAAGAATCG		GCGACGGCTA	CGCCAAAGTG	ATGAGCGACG
701	AATTTATCGC		AAGCTTTTTG	CCGAGCAGGC	
751	GACATCATCA		CGCCATTCCG	GGCAAACCCG	41-41010
801	NNTNANCAAA		AAAGCATGAA	ACCCGGCTCC	GTCATCGTCG
851	ATTTGGCGGC	GGCGACGGGC	GGCAACTGCG		ACAGGGCGAA
901	TTGTTCGTAA	CCGGCAACGG	CGTGAAAATC	ATCGGCTACA	
951	AAACCGCCTT	GCCGGACAGT	CTTCGCAGCT	TTACCCCAACA	CCGACATGGC
1001	ACCTGACCAA		CCGAACAAAG		AACTTGGTCA
1051	TTCGAAGACG	TGATTATCCG	CAACATGACC	ACGGCGAAAT	
1101	CACCTTCCCG		TTCAAGTTTC	GTTACCCGCG	ACGGCGAAAT
1151	CGTCTGAAAA		GCCGCCAAGC	CGCCCAACCG	CAGCAAACGC
		30000001	GCCGCCAAGC	CCGAACCGAA	A CCCCTTCCC

1201 1251 1301 1351 1401 1451 1501	TCCTCGCCTG CGTCATCGGC TACTATGTCG TTTGGAACGT CAGCCACTCG CTGCACACAC CGCTGATGTC GGTGACCAAC GCCATTTCCG GCATCATCGT CGTCGGCGCG CTGCTGCAAA TCGGTCAGGG CAACGGCTTC GTTTCGCTGC TGTCGTTTGT TGCCATCCTG ATTGCCAGCA TCAACATCTT CGGCGGCTTC TTTGTAACGC GGCGGATGCT GAATATGTTT AGGAAAGGGT AA	
This correspond	ds to the amino acid sequence <seq 155.a="" 642;="" id="" orf="">:</seq>	
1	MKIGIPRESL SGETRVACTP ATVALLGKLG FETVVESGAG LAASLDDAAY	
51 101	QAAGATVADK AAVWAYPLIY KVNAPSEDEL PLLKEGQTIV SFLWPRQNEA LVEALRAKKV NALAMDMVPR ISRAQALDXL SXMANISGYR AVIEAANAFG	
151	RXFTGQITAA GKVPPAQVLV IGAGVAGLAA IGTANSLGAV VRVFDTRLXV	
201 251	AEQLESMGGK FLKLDFPQES GGSGDGYAKV MSDEFIAAEM KLFAEQAKEV DIIITTAAIP GKPAPKXXXK EMVESMKPGS VIVDLAAATG GNCELTKQGE	
301	LFVTGNGVKI IGYTDMANRL AGQSSQLYAT NLVNLTKLLS PNKDGEITLD	
351	FEDVIIRNMT VTRDGEITFP PPPIQVSAQP QQTPSEKAAP AAKPEPKPVP	
401 451	LWKK <u>LAPAXI AAVLVLWVGA V</u> APAAFLNHF IVFVLACVIG YYVVWNVSHS LHTP <u>LMSVTN AISGIIVVGA L</u> LQIGQGNGF <u>VSLLSFVAIL</u> IASINIFGGF	
501	FVTRRMLNMF RKG*	
m155/a155	95.3% identity in 513 aa overlap	
4.5.5	10 20 30 40 50	50
m155.pep	MKIGIPRESLSGETRVACTPATVALLGKLGFETVVESGAGLAASLDDAAYQTAGATVAI	)K
a155	MKIGIPRESLSGETRVACTPATVALLGKLGFETVVESGAGLAASLDDAAYQAAGATVAI	K
	10 20 30 40 50	50
	70 80 90 100 110 12	
m155.pep	AAVWVCPLIYKVNAPSEQELPLLNEGQTIVSFLWPRQNEALVEALRAKKVNALAMDMVI     :           :    :	PR:
a155	AAVWAYPLIYKVNAPSEDELPLLKEGQTIVSFLWPRQNEALVEALRAKKVNALAMDMVI	PR
	70 80 90 100 110 12	20
		30
m155.pep	ISRAQALDALSSMAN   SGYRAV   EAANAFGRFFTGQ   TAAGKVPPAQVLV   GAGVAGLA 	
a155	ISRAQALDXLSXMANISGYRAVIEAANAFGRXFTGQITAAGKVPPAQVLVIGAGVAGLA	ι. LA
	130 140 150 160 170 18	10
	190 200 210 220 230 24	
m155.pep	IGTANSLGAVVRAFDTRLEVAEQIESMGGKFLKLDFPQESGGSGDGYAKVMSDEFIAAF 	M
a155	IGTANSLGAVVRVFDTRLXVAEQLESMGGKFLKLDFPQESGGSGDGYAKVMSDEFIAAE	
	190 200 210 220 230 24	0
	250 260 270 280 290 30	
m155.pep	KLFAEQAKEVDIIITTAAIPGKPAPKLITKEMVESMKSGSVIVDLAAATGGNCELTRPG	iΕ
a155	KLFAEQAKEVDIIITTAAIPGKPAPKXXXKEMVESMKPGSVIVDLAAATGGNCELTKQC	ξE
	250 260 270 280 290 30	0 (
m155.pep	310 320 330 340 350 36	
mrss.pep	LSVTGNGVKIIGYTDMANRLAGQSSQLYATNLVNLTKLLSPNKDGEITLDFEDVIIRNN 	11.
<b>a15</b> 5	LFVTGNGVKIIGYTDMANRLAGQSSQLYATNLVNLTKLLSPNKDGEITLDFEDVIIRNN	1T
	310 320 330 340 350 36	0
m155 non	370 380 390 400 410 42	0:
m155.pep	VTHDGEITFPPPPIQVSAQPQQTPSEKAVPAAKPEPKPVPLWKKLAPAVIAAVLVLWVG	iA 
a155	VTRDGEITFPPPPIQVSAQPQQTPSEKAAPAAKPEPKPVPLWKKLAPAXIAAVLVLWVQ	A
	370 380 390 400 410 42	0

BNSDOCID: <WO___9957280A2_I_>

m155.pep	430 VAPAAFLNHFIVFV	440 LACVIGYYVV	450 WNVSHSLHTI	460 PLMSVTNAISGI	470 IVVGALLO	480 IGQGNGF		
a155	VAPAAFLNHFIVFV	΄''''ΙΙΙΙΙΙΙΙΙΙΙΙΙΙΙΙΙΙΙΙΙΙΙΙΙΙΙΙΙΙΙΙΙ			1	111111		
	400	Line Vigit VV	MMADUDTUIT	PLMSVTNAISGI	IVVGALLQ	IGQGNGF		
	430	440	450	460	470	480		
	490	500	510					
m155.pep	VSLLSFVAILIAGI		RMLNMFKKGX	ζ . <del>.</del>				
	-		111111:11					
a155	VSLLSFVAILIASINIFGGFFVTRRMLNMFRKGX							
	490	500	510 .	•				

```
The following partial DNA sequence was identified in N. gonorrhoeae <SEQ ID 643>:
```

```
g156.seq
               ATGACTTTCG CCTATTGGTG CATTCTGATT GCCTGCCTAT TGCCGCTTTT
           51
               TTGTGCGGCG TATGCCAAAA AAGCGGGCGG ATTCCGGTTT AAAGACAACC
          101 ACAATCCTCG CGGTTTTCTG GCACATACGC AAGGCGCAGC CGCCCGTGCC
          151 CACGCCGCG AGCAAAACGG TTTTGAAGCC TTTGCACCGT TTGCCGCCGC
          201 CGTTTTGACG GCACACGCAA CCGGCAATGC CGGACAAGCA ACCGTCAACA
              CGCTTGCCGG ATTGTTCATC CTGTTCCGCC TCGCCTTTAT CTGGTGCTAC
          301 ATCGCAGACA AAGCAGCATT GCGCTCGCTG ATGTGGGCGG GCGGATTTGC
          351 CTGCACCGTC GGACTGTTTG TCGCGGCTGC TTGA
This corresponds to the amino acid sequence <SEQ ID 644; ORF 156.ng>:
```

g156.pep

MTFAYWCILI ACLLPLFCAA YAKKAGGFRF KDNHNPRGFL AHTQGAAARA HAAQQNGFEA FAPFAAAVLT AHATGNAGQA TVNTLAGLFI LFRLAFIWCY 1

IADKAALRSL MWAGGFACTV GLFVAAA*

#### The following partial DNA sequence was identified in N. meningitidis <SEQ ID 645>: m156.seq

ATGACTTTCG CCTATTGGTG TATTCTGATT GCCTGCCTAT TGCCGCTTTT TTGTGCGGCG TATGCCAAAA AAGCGGGCGG ATTCCGGTTT AAAGACAACC 51 101 ACAATCCGCG CGGTTTTCTA GCGCACACGC AAGGCGCAGC CGCCCGTGCC 151 CACGCCGCAC AGCAAAACGG TTTTGAAGCC TTTGCACCGT TTGCCGCCGC 201 CGTTTTGACG GCACACGCAA CCGGCAATGC GGCGCAATCG ACCATCAACA 251 CGCTTGCCTG CCTGTTCATC CTGTTCCGCC TCGCCTTTAT CTGGTGCTAT
301 ATCGCCGACA AAGCCGCTAT GCGCTCACTG ATGTGGGCAG GCGGATTTGC 351 CTGCACCGTC GGGCTGTTTG TCGCGGCTGC TTGA

#### This corresponds to the amino acid sequence <SEQ ID 646; ORF 156>:

m156.pep

MTFAYWCILI ACLLPLFCAA YAKKAGGFRF KDNHNPRGFL AHTQGAAARA 1 HAAQQNGFEA FAPFAAAVLT AHATGNAAQS TINTLACLFI LFRLAFIWCY 51

IADKAAMRSL MWAGGFACTV GLFVAAA*

#### Computer analysis of this amino acid sequence gave the following results: Homology with a predicted ORF from N. gonorrhoeae

m156 / g156 96.1% identity in 127 aa overlap

m156.pep	10	20	30	40	50	60		
mroo.pep	MTFAYWCILIACLLPLFCAAYAKKAGGFRFKDNHNPRGFLAHTQGAAARAHAAQQNGFEA							
q156		111111111	111111111		1111111111			
3	MTFAYWCILIACLLPLFCAAYAKKAGGFRFKDNHNPRGFLAHTQGAAARAHAAQQNGFEA							
	10	20	30	40	50	60		
m156.pep	70	80	90	100	110	120		
тестрор	FAPFAAAVLTAHATG	NAAQSTINT.	LACLFILFRI	AFIWCYIADK	AAMRSLMWAG	GFACTV		
	11111111111111	11:1:1:11	11 111111	$\Pi\Pi\Pi\Pi\Pi\Pi$	11:111111	111111		